Environmental Perturbations Lift the Degeneracy of the Genetic Code to Regulate Protein Levels in Bacteria

Environmental perturbations lift the degeneracy of the genetic code to regulate protein levels in bacteria

Arvind R. Subramaniam[a], Tao Pan[b], and Philippe Cluzel[a]

[a] FAS Center for Systems Biology, Department of Molecular and Cellular Biology, and School of Engineering and Applied Sciences, Harvard University, 52 Oxford St, Cambridge, MA 02138, USA.

[b] Department of Biochemistry and Molecular Biology, and Institute of Biophysical Dynamics, University of Chicago, Chicago, IL 60637, USA.

Correspondence and requests for materials should be addressed to P.C. (cluzel@mcb.harvard.edu).

Classification: Biological Sciences - Systems Biology.

Abstract

The genetic code underlying protein synthesis is a canonical example of a degenerate biological system. Degeneracies in physical and biological systems can be lifted by external perturbations thus allowing degenerate systems to exhibit a wide range of behaviors. Here we show that the degeneracy of the genetic code is lifted by environmental perturbations to regulate protein levels in living cells. By measuring protein synthesis rates from a synthetic reporter library in *Escherichia coli*, we find that environmental perturbations, such as reduction of cognate amino acid supply, lift the degeneracy of the genetic code by splitting codon families into a hierarchy of robust and sensitive synonymous codons. Rates of protein synthesis associated with robust codons are up to 100-fold higher than those associated with sensitive codons under these conditions. We find that the observed hierarchy between synonymous codons is not determined by usual rules associated with tRNA abundance and codon usage. Rather, competition among tRNA isoacceptors for aminoacylation underlies the robustness of protein synthesis. Remarkably, the hierarchy established using the synthetic library also explains the measured robustness of synthesis for endogenous proteins in *E. coli*. We further found that the same hierarchy is reflected in the fitness cost of synonymous mutations in amino acid biosynthesis genes and in the transcriptional control of sigma factor genes. Our study reveals that the degeneracy of the genetic code can be lifted by environmental perturbations, and it suggests that organisms can exploit degeneracy lifting as a general strategy to adapt protein synthesis to their environment.

\body

Introduction

Degeneracy, the occurrence of distinct states that share a common function, is a ubiquitous property of physical and biological systems (1–3). Examples of degenerate systems include atomic spectra (4), condensed matter (5), the nervous system (2) and the genetic code (6, 7). Degeneracy in physical systems is often associated with underlying symmetries (1), and in biological systems with error-minimization, evolvability, and robustness against perturbations (8). Degenerate states that are indistinguishable under normal conditions can exhibit distinct properties under the action of external perturbations (1). This effect, called degeneracy lifting, allows degenerate systems to exhibit a wide range of behaviors depending on the environmental context (2). The genetic code governing protein synthesis is a highly degenerate system since 18 of the 20 amino acids have multiple synonymous codons and 10 of the 20 amino acids are aminoacylated (charged) onto multiple tRNA isoacceptors. Protein synthesis rates in living cells respond to diverse environmental perturbations, which raises the question of whether any of these perturbations modulates protein levels by lifting the degeneracy of the genetic code. Previous experiments found that both the concentration of charged tRNAs as well as the occupancy of ribosomes on synonymous codons undergo significant changes upon nutrient limitation (9–11). Yet whether such environmental perturbations lift the degeneracy of the genetic code by modulating the expression level of proteins is unknown. Here, we propose to use amino acid limitation in the bacterium *Escherichia coli* as a model system to investigate whether the degeneracy of the genetic code can be lifted by environmental perturbations, and how

degeneracy lifting could provide a general strategy to adapt protein synthesis to environmental changes.

Results

Degeneracy lifting upon amino acid limitation

We considered synonymous codons for seven amino acids: Leu, Arg, Ser, Pro, Ile, Gln, and Phe. This set of seven amino acids is representative of the degeneracy of the genetic code, in that it includes six-, four-, three- and two-fold degenerate codon families. We constructed a library of 29 *yellow fluorescent protein* (*yfp*) gene variants, each of which had between six and eight synonymous mutations for one of the seven amino acids (Fig. 1A). In this library, we designed each *yfp* variant to characterize the effect of one specific codon on protein synthesis. We expressed the *yfp* variants constitutively at low gene dosage (2 copies / chromosome, Fig. 1B) in *E. coli* strains that were auxotrophic for one or more amino acids. We monitored growth and YFP synthesis in these strains during amino acid-rich growth as well as during limitation for each of the seven amino acids (Methods).

During amino acid-rich growth, our measurements revealed that protein synthesis rates were highly similar across *yfp* variants, with less than 1.4-fold variation within all codon families (Fig. 1D, grey bars). Thus, under rich conditions, the degeneracy of the genetic code remains intact with respect to protein synthesis. Strikingly, under amino acid-limited growth, codon families split into a hierarchy of YFP synthesis rates (Fig. 1C, 1D). We found that some synonymous codons, such as CTA for leucine, were highly sensitive to environmental perturbation, causing YFP synthesis rates to be near zero in response to the limitation of these

codons' cognate amino acids. Conversely, other synonymous codons, such as CTG for leucine, were more robust to the same perturbation with synthesis rates of YFP up to 100-fold higher than the sensitive ones[*]. In addition to fluorescence, this difference in robustness was reflected in protein levels measured with Western blotting (Fig. S1). Notably, even a single substitution to a perturbation-sensitive codon in the *yfp* coding sequence resulted in more than a 2-fold difference in YFP synthesis rate during limitation for the cognate amino acid, without any effect on synthesis rate during amino acid-rich growth (Fig. S2). Only those codons that were cognate to the limiting amino acid caused splitting of YFP synthesis rates (Fig. S3). Interestingly, the splitting was more acute for codon families with six-fold degeneracy (Leu, Arg, Ser), while splitting was weaker for codon families with four-, three- and two-fold degeneracies (Fig. 1D, first row vs. second row). These results support the idea that greater degeneracy typically allows systems to exhibit wider range of responses to environmental perturbations (2). In subsequent experiments, we focused on the two codon families, leucine and arginine that displayed the largest range of splitting. These two families constitute 16% of codons across the genome of *E. coli*.

Intracellular determinants of the hierarchy among synonymous codons

We sought to identify the intracellular parameters that determine the observed hierarchy of degeneracy splitting during amino acid limitation. To this end, we quantified the robustness of synthesis rate to amino acid limitation as the ratio of YFP synthesis rates between amino acid-limited and amino acid-rich growth phases. Protein synthesis rate is known to be correlated with

[*] We define codons as robust when the synthesis rate from the corresponding *yfp* variant during cognate amino acid limitation is higher than the average synthesis rate within that codon family. Similarly, we define codons as sensitive when the synthesis rate from the corresponding *yfp* variant during cognate amino acid limitation is lower than the average synthesis rate within that codon family.

codon usage and tRNA abundance during artificial over-expression of proteins (12, 13). However, we found that robustness of YFP synthesis to amino acid limitation was not correlated with either codon usage or tRNA abundance (r^2 = 0.08 and 0.00, squared Spearman rank-correlation, Fig. S4). We then considered determinants of protein synthesis that might be important specifically during amino acid limitation. tRNA isoacceptors are uniformly charged (aminoacylated) at about 80% under amino acid-rich conditions (14, 15). However during perturbations such as amino acid limitation, some tRNA isoacceptors cognate to the amino acid are almost fully charged while other isoacceptors in the same family have charged fractions that are close to zero (10, 16). A theoretical model proposed that such selective charging arises from differences in the relative supply and demand for charged tRNA isoacceptors (9). While it is unclear how this mechanism could solely control protein levels, charged tRNA play an essential role as substrates for the elongation of ribosomes across individual codons (17). Consequently, we hypothesized that selective charging of tRNA isoacceptors also underlies the observed splitting in synthesis rates among *yfp* variants. Consistent with this hypothesis, charged fractions of leucine and arginine tRNA isoacceptors during limitation of cognate amino acid starvation measured in a previous work (10) were correlated with the robustness of synthesis rates from *yfp* variants after accounting for codon-tRNA assignments (r^2 = 0.78, Fig. S5).

We experimentally tested whether varying the concentration of charged tRNA could change the hierarchy of protein synthesis rates initially revealed by amino acid limitation. To this end, we co-expressed each one of the leucine or arginine tRNA isoacceptors together with each of the six leucine or arginine variants of *yfp*, respectively (Fig. 2). Previous work (16) showed that overexpression of a single tRNA isoacceptor cognate to a limiting amino acid enables it to compete better in the common charging reaction against other isoacceptors. As a result, charged

tRNA concentration of the overexpressed isoacceptor increases, while charged tRNA concentrations of the remaining isoacceptors for that amino acid decrease or remain unchanged (16). We found that *yfp* variants constructed with perturbation-sensitive codons exhibited higher synthesis rates upon co-expression of tRNA isoacceptors cognate to those perturbation-sensitive codons (Fig. 2, A-B , bottom three rows, solid black-outlined squares). Conversely, *yfp* variants with perturbation-robust codons exhibited lower protein synthesis rates upon co-expression of non-cognate tRNA isoacceptors (Fig. 2, A-B, top three rows, non-outlined squares). These two patterns of changes in YFP synthesis rate mirror previously measured changes in charged tRNA concentration upon tRNA co-expression (16), thereby suggesting that the observed hierarchy in synthesis rates of *yfp* variants are tightly coupled with the concentrations of cognate charged tRNA isoacceptors during amino acid limitation. By contrast, tRNA co-expression did not affect synthesis rates from *yfp* variants in the absence of perturbation, i.e., during amino acid-rich growth (Fig. 2C). We observed several codon-tRNA pairs with mismatches at the wobble position but that do not satisfy known wobble-pairing rules (Table S9), and that showed an increase in YFP synthesis rate upon co-expression of the tRNA isoacceptor during amino acid limitation (Fig. 2, A-B , dashed black-outlined squares).

A codon robustness index for endogenous proteins

We investigated whether the hierarchy of synthesis rates measured for the synthetic *yfp* variants also governs the synthesis of endogenous proteins of *E. coli*. We first devised a general parameter, hereafter called the codon robustness index (CRI), to characterize the robustness of any protein's synthesis rate to an environmental perturbation associated with limitation of a specific amino acid (Fig. 3A). We defined CRI as a product of codon-specific weights w_{codon}, and we inferred these weights from the synthesis robustness of *yfp* variants to limitation for their

cognate amino acid (Fig. 3B). Our formulation of CRI is based on the simplifying assumption that each codon decreases protein synthesis rate by a factor w_{codon} that is independent of the codon's intragenic location, the presence of other codons in the coding sequence, or the specific cellular role of the encoded protein. By definition, w_{codon} is unity for codons that are not cognate to the limiting amino acid, and perturbation-robust codons have a higher w_{codon} value than perturbation-sensitive codons for the limiting amino acid.

To test the predictive power of CRI, we selected 92 *E. coli* open reading frames (ORFs) that span a broad range of leucine CRI values and functional categories (Fig. S7, Table S1). We expressed the corresponding proteins constitutively as N-terminus fusions with YFP[†] in an *E. coli* strain auxotrophic for leucine (Fig. 3C, Inset,). Upon leucine limitation, we found a strong correlation between the robustness of protein synthesis rates from the 92 ORF-*yfp* fusions and their leucine CRI values (Fig. 3C, r^2=0.61, P=10^{-23}, squared Spearman rank-correlation). Similarly, arginine CRI was also strongly correlated with robustness of a library of 56 ORF-*yfp* fusions during arginine limitation (r^2=0.59, P=10^{-12}, Fig. S8, Table S2). By contrast, standard measures of translation efficiency under amino acid-rich conditions such as codon adaptation index (18), tRNA adaptation index (19) or folding energy of the mRNA around the start codon (20) displayed only a weak correlation with protein synthesis rate from the ORF-*yfp* fusions during amino acid-rich growth (r^2 = 0.10, 0.08, and 0.02 resp., Fig. S9). We further found that changes in Leu CRI calculated from the *yfp* data could predict both the effect of tRNA co-expression and that of synonymous mutations on protein synthesis from *E. coli* ORFs during

[†] The YFP fusion partner in the 92 ORF-yfp fusions used for testing Leu CRI was encoded by the CTG variant of *yfp* that has the highest, most robust synthesis rate during leucine limitation. Similarly, the YFP fusion partner in the 56 ORF-*yfp* fusions used for testing Arg CRI was encoded by the AGA variant of *yfp* that has the highest, most robust synthesis rate during arginine limitation.

leucine limitation (Fig. 3D, Fig. S10). Importantly, similar to our results using *yfp* reporters, neither tRNA co-expression nor synonymous mutations for *E.coli* ORF-*yfp* fusions had a significant effect on the synthesis rates from these ORFs during leucine-rich growth in absence of environmental perturbations (Fig. S11). Thus the degeneracy of the genetic code underlies the levels of endogenous protein production only during response to environmental perturbations.

Consequences of degeneracy lifting for fitness and gene regulation

Degeneracy splitting in physical systems can be exploited to encode information related to the environmental context (21, 22). We asked whether bacteria might similarly exploit the degeneracy splitting of genetic code during response to amino acid limitation. Hence we tested whether the expression of amino acid biosynthesis genes that enable bacteria to adapt to amino acid limitation is affected by the hierarchy between robust and sensitive codons. We found that mutating codons that are perturbation-robust to those that are perturbation-sensitive in the leucine-biosynthesis genes *leuA*, *leuC* and *leuD*, and the arginine biosynthetic gene *carA* decreased their protein synthesis rate during cognate amino acid limitation, but not during amino acid-rich growth (Fig. S12). Interestingly, in the case of *leuA* and *carA*, the same synonymous mutations also resulted in a fitness cost for prototrophic strains upon downshift from amino acid-rich to amino acid-poor conditions (Fig. 4A). Thus synonymous mutations can have a significant fitness cost during an environmental perturbation, which is distinct from that measured under nutrient-rich conditions in the absence of any perturbation (20, 23). However swapping codons that are perturbation-robust to those that are perturbation-sensitive in other biosynthesis genes (see *argA* and *leuC* in Fig. 4A) did not significantly affect fitness, suggesting that the hierarchy of robust and sensitive codons might be selectively utilized by bacteria to regulate genes within a single metabolic pathway.

Perturbations associated with amino acid limitation in *E. coli* can result in two distinct outcomes, depending on the environmental conditions: On one hand, when substrates used in amino acid biosynthesis are still abundant in the environment, the cell up-regulates corresponding biosynthesis genes to mitigate the limitation of amino acids and resume growth. On the other hand, in the absence of substrates for amino acid biosynthesis, *E. coli* can survive a prolonged period in amino acid-poor environments through a cellular response mediated by sigma factors (24, 25). We found that genes encoding several stress-response sigma factors (*rpoS*, *rpoE* and *rpoH*) are enriched in TTA and TTG, the leucine codons that ensure robust protein synthesis during leucine limitation (Fig. 4B, top panel). By contrast, genes for the housekeeping sigma factor (*rpoD*) and a few minor sigma factors (*fecI*, *fliA*) are enriched for CTC and CTT, which are sensitive to leucine limitation. This contrasting pattern is observed for leucine (but not for arginine), and is further mirrored by the change in transcript abundance for sigma factor genes in response to leucine limitation (Fig. 4B, bottom panel). Hence degeneracy splitting in the genetic code might be exploited in concert with transcriptional control to regulate protein levels.

Discussion

In summary, we have found that the degeneracy of the genetic code does not have a role in regulating protein synthesis during amino acid-rich growth. By contrast, the splitting of this degeneracy upon reduction in amino acid supply has a potent effect on protein synthesis that results in up to 100-fold differences in protein synthesis rates between synonymous gene variants. Such a large role for synonymous codons in protein synthesis is surprising given that other post-transcriptional mechanisms such as protein degradation are known to play a

significant role upon amino acid limitation (26). We identified competition between tRNA isoacceptors for aminoacylation as a key determinant of the hierarchy of protein synthesis rates during amino acid limitation. Low concentration of a charged tRNA isoacceptor can cause ribosomes to selectively pause at its cognate codon[‡] and trigger ribosome jamming (27), translation-recoding (28), mRNA cleavage (29–31) or feedback-transcriptional control (32, 33). It will be interesting to find the relative contribution of these different molecular processes to the degeneracy lifting uncovered here[§]. Here, we have investigated the effect of a specific environmental perturbation associated with amino acid limitation in the bacterium *E. coli*. However, this type of perturbation plays a crucial role in the lifecycle of other bacteria such as *Myxococcus xanthus* and *Bacillus subtilis* that undergo differentiation cued by amino acid limitation (34, 35). Protein synthesis during such differentiation events might also be regulated by degeneracy lifting of the genetic code. Moreover, degeneracy lifting could be important during protein synthesis in eukaryotes, where clinically-important conditions such as neoplastic transformation and drug treatment are often accompanied by a reduction in amino acid supply (36, 37). Therefore lifting the degeneracy of the genetic code might emerge as a general strategy for biological systems to expand their repertoire of responses to environmental perturbations.

[‡] A recent genome-wide study found increased ribosome pausing at serine codons during serine-limited growth of *E. coli*. Interestingly, ribosomes paused significantly only at four out of the six serine codons, and these four codons are precisely the same ones that caused YFP synthesis rate to be sensitive to serine limitation in our experiments (Fig. S13).

[§] We measured the change in mRNA levels of different *yfp* variants in response to amino acid limitation. Changes in mRNA levels were correlated with corresponding changes in YFP synthesis rates upon amino acid limitation (Fig. S14). However, changes in mRNA levels were smaller than expected, suggesting that changes in mRNA abundance induced by ribosome pausing might not be solely responsible for the observed changes in protein synthesis rate.

Materials and Methods

Summary of key methods are given below. Detailed methods for all experiments and analyses are included in Appendix.

Bacterial strains

All strains used in this study were obtained from the *E.coli* Genetic Stock Center (CGSC), Yale University. Different auxotrophic strains were used depending on the amino acid that was limiting in the growth medium (Table S5). Strains were stored as 20% glycerol stocks at -80C either in 1ml cryo-vials or in 96-well plates (3799, Costar). For experiments involving over 25 strains, a temporary 20% glycerol stock was stored at -20C in 96-well PCR plates.

Plasmids

The *pZ* series of plasmids (39) were used for expression of all genes constructed for this study. General features of the plasmid backbones are described here. Details on individual gene constructs that were inserted into these plasmid backbones, including DNA sequences and plasmid maps, are in Appendix. A low-copy plasmid, *pZS*11* [*SC101* ori* (3-4 copies/cell), Amp^R (*bla* gene) and a constitutive $P_{LtetO-1}$ promoter] was used for expression of all fluorescent reporter genes and their fusions. The synthetic ribosome binding site (RBS) in the original *pZS*11* backbone was replaced by a modified *T7*-based RBS that resulted in efficient expression of most coding sequences. A medium-copy plasmid, *pZA32* [*p15A ori* (10-12 copies/cell), Chl^R (*cat* gene) and $P_{LlacO-1}$ promoter] was used for expression of all tRNA genes. Strains with *pZA32* plasmids were grown with 1mM IPTG to ensure constitutive expression of all tRNA genes. Standard plasmids *pUC18* and *pUC19* were used as intermediate cloning vectors for site-directed mutagenesis.

Gene synthesis and cloning

A single *yfp* sequence was built *de novo* (synthesis by Genscript, USA). All subsequent *yfp* variants were constructed using a site-directed mutagenesis kit (Stratagene). tRNA genes and *E. coli* ORFs were amplified from the chromosome of a wild-type *E. coli* strain (MG1655) by PCR (Details on cloning and genes sequences in Appendix).

Amino acid limitation experiments

Overnight cultures were inoculated from glycerol stocks or fresh colonies and grown in a MOPS-based rich-defined medium with 800µM of 19 amino acids and 10mM serine at 30C with shaking. For experiments involving amino acid limitation, overnight cultures were diluted 1:1000 into a similar rich-defined medium as the overnight cultures. However the amino acid, whose limitation was to be induced, was added at a reduced concentration and supplemented with its methyl-ester analog (see Table S6 for exact concentrations). Amino acid methyl-esters are inefficiently metabolized analogs of the corresponding amino acids and have been previously used for steady growth of *E. coli* under amino acid limiting conditions (40, 41) (see Figs. S15 and S16 for the effect of methyl-ester on growth and measured robustness during amino acid

limited growth). Slight variations in the initial concentration of either the limiting amino acid or its methyl-ester only results in shifting of the transition to a higher or lower cell density without appreciable changes in growth rate (see Notes S1 and S2). Growth and fluorescence were quantified using a standard 96-well plate reader integrated with a robotic system. Further details on growth protocols are given in Appendix.

Analysis of cell density and fluorescence time series

Matlab R2009 (Mathworks) was used for all analyses unless otherwise mentioned. All correlations and P-values reported in this work were calculated using the Matlab command 'corr' with the 'Type' option set to either 'Spearman' or 'Pearson' as appropriate. Growth and fluorescence time series were fit with exponential and linear curves in the amino acid rich and amino acid limited growth regimes, respectively, and the onset time of amino acid limited growth was automatically inferred from their intersection. Protein synthesis rate, S was calculated as:

$$\text{Protein synthesis rate}, S = \frac{1}{\text{Absorbance}} \times \frac{d\ (\text{Fluorescence})}{d\ (\text{time})}$$

First, the above formula was evaluated at the onset time of amino acid limited growth using the exponential fits for absorbance and fluorescence data in the amino acid rich growth regime. Next, the same formula was evaluated at the onset time using the linear fits for absorbance and fluorescence data in the amino acid limited growth regime. These two values correspond to the protein synthesis rates reported for the amino acid rich and amino acid limited growth regimes (such as the data in Fig. 1D). Further details of this analysis are given in Appendix.

Calculation of CRI

CRI for a protein coding sequence corresponding to a limiting amino acid was calculated by multiplying the w_i values for codons cognate to the limiting amino acid in that sequence. w_i values shown in Fig. 3B were calculated using the robustness of protein synthesis of the corresponding *yfp* variants during cognate amino acid limitation (Fig. 1D). Based on our non-cognate limitation experiment (Fig. S2), the w_i values for all codons other than those cognate to the limiting amino acid are set to be equal to 1.

For illustration, we demonstrate the calculation of w_i for the 6 Leu codons below. The exact same procedure was followed for other synonymous codon families. Taking $\log_2 w_i \equiv W_i$ for each codon, and $\log_2$ robustness during amino acid limited growth $\equiv SR$ for each *yfp* variant,

$$7 \times W_{CTA} + 15 \times W_{CTG} = SR_{yfp,CTA}$$

$$7 \times W_{CTC} + 15 \times W_{CTG} = SR_{yfp,CTC}$$

$$22 \times W_{CTG} = SR_{yfp,CTG}$$

$$7 \times W_{CTT} + 15 \times W_{CTG} = SR_{yfp,CTT}$$

$$7 \times W_{TTA} + 15 \times W_{CTG} = SR_{yfp,TTA}$$

$$7 \times W_{TTG} + 15 \times W_{CTG} = SR_{yfp,TTG}$$

The multiplicative factors on the LHS in front of W_i correspond to the number of different Leu codons in the corresponding Leu variant of *yfp* (see Fig. 1A). The RHS is the measured ($\log_2$) robustness of protein synthesis from the corresponding *yfp* variant during Leu limitation (see Fig. 1D). These equations were solved simultaneously to determine the w_i value for each Leu codon. Revised w_i values based on *yfp* measurements in the presence of ^{GAG}Leu2 tRNA (Fig. 2) were used for calculation of Leu CRI in the case of ^{GAG}Leu2 tRNA co-expression with *E. coli* ORFs (Fig. 3D).

Acknowledgements

We acknowledge J. Shapiro for work on developing the preliminary growth assays, and J. Gallant and M. Cashel for useful correspondence on optimizing the growth assays. A.S. is grateful to C. C. Guet for help with cloning and suggesting the tRNA co-expression experiments, and J. Moffitt for pointing him to the CP78 *E. coli* strain. We thank B. Stern, A. Murray, and V. Denic for critical questions. We thank K. Dave for editorial advice and assistance, and the members of the Cluzel lab and FAS Center for Systems Biology for experimental support. We acknowledge M. Aldana, L. David, D. A. Drummond, J. Elf, M. Kim, L. Marshall, M. Mueller, E. O'Shea, E. Wallace, J. Weissman, K. Wood, and B. Zid for comments on earlier versions of the manuscript. We are grateful to the anonymous referees for critical comments on the tRNA co-expression experiment. Author Contributions: A.S., T.P. and P.C. designed research. A.S. performed all experiments. A.S, T.P. and P.C. analyzed data and wrote the manuscript.

References

1. Shankar R (1994) *Principles of quantum mechanics* (Plenum Press, New York). 2nd Ed.
2. Edelman GM, Gally JA (2001) Degeneracy and Complexity in Biological Systems. *PNAS* 98:13763–13768.
3. Baer M, Billing GD (2002) *The role of degenerate states in chemistry* (J. Wiley & Sons, Hoboken, N.J.).

4. Cowan RD (1981) *The theory of atomic structure and spectra* (University of California Press, Berkeley).

5. Affleck I, Ludwig AWW (1991) Universal noninteger "'ground-state degeneracy'" in critical quantum systems. *Phys Rev Lett* 67:161–164.

6. Crick F (1955) *On Degenerate Templates and the Adaptor Hypothesis: A Note for the RNA Tie Club*.

7. Reichmann ME, Rees MW, Symons RH, Markham R (1962) Experimental Evidence for the Degeneracy of the Nucleotide Triplet Code. 195:999–1000.

8. Stelling J, Sauer U, Szallasi Z, Doyle III FJ, Doyle J (2004) Robustness of Cellular Functions. *Cell* 118:675–685.

9. Elf J, Nilsson D, Tenson T, Ehrenberg M (2003) Selective charging of tRNA isoacceptors explains patterns of codon usage. *Science* 300:1718–22.

10. Dittmar KA, Sorensen MA, Elf J, Ehrenberg M, Pan T (2005) Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO Rep* 6:151–157.

11. Li GW, Oh E, Weissman JS (2012) The anti-Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* 484:538–541.

12. Robinson M et al. (1984) Codon usage can affect efficiency of translation of genes in Escherichia coli. *Nucleic Acids Res* 12:6663–6671.

13. Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. *Trends Biotechnol* 22:346–353.

14. Krüger MK, Sorensen MA (1998) Aminoacylation of hypomodified tRNAGlu in vivo. *J Mol Biol* 284:609–620.

15. Sorensen MA (2001) Charging levels of four tRNA species in Escherichia coli Rel(+) and Rel(-) strains during amino acid starvation: a simple model for the effect of ppGpp on translational accuracy. *J Mol Biol* 307:785–98.

16. Sorensen MA et al. (2005) Over expression of a tRNA(Leu) isoacceptor changes charging pattern of leucine tRNAs and reveals new codon reading. *J Mol Biol* 354:16–24.

17. Ibba M, Söll and D (1999) Quality Control Mechanisms During Translation. *Science* 286:1893–1897.

18. Sharp PM, Li WH (1987) The Codon Adaptation Index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* 15:1281–1295.

19. dos Reis M, Savva R, Wernisch L (2004) Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res* 32:5036–5044.

20. Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-Sequence Determinants of Gene Expression in Escherichia Coli. *Science* 324:255–258.

21. Gershenfeld NA, Chuang IL (1997) Bulk Spin-Resonance Quantum Computation. *Science* 275:350–356.

22. Chuang IL, Vandersypen LMK, Zhou X, Leung DW, Lloyd S (1998) Experimental realization of a quantum algorithm. *Nature* 393:143–146.

23. Lind PA, Berg OG, Andersson DI (2010) Mutational Robustness of Ribosomal Protein Genes. *Science* 330:825–827.

24. Traxler MF et al. (2011) Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the "feast to famine" gradient in Escherichia coli. *Mol Microbiol* 79:830–845.

25. Durfee T, Hansen A-M, Zhi H, Blattner FR, Jin DJ (2008) Transcription profiling of the stringent response in Escherichia coli. *J Bacteriol* 190:1084–96.

26. Kuroda A et al. (2001) Role of Inorganic Polyphosphate in Promoting Ribosomal Protein Degradation by the Lon Protease in E. coli. *Science* 293:705–708.

27. Tuller T et al. (2010) An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* 141:344–54.

28. Zaher HS, Green R (2011) A Primary Role for Release Factor 3 in Quality Control during Translation Elongation in Escherichia coli. *Cell* 147:396–408.

29. Christensen SK, Gerdes K (2003) RelE toxins from Bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. *Mol Microbiol* 48:1389–1400.

30. Li X, Yagi M, Morita T, Aiba H (2008) Cleavage of mRNAs and role of tmRNA system under amino acid starvation in Escherichia coli. *Mol Microbiol* 68:462–473.

31. Garza-Sánchez F, Gin JG, Hayes CS (2008) Amino Acid Starvation and Colicin D Treatment Induce A-site mRNA Cleavage in Escherichia coli. *J Mol Biol* 378:505–519.

32. Henkin TM, Yanofsky C (2002) Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays* 24:700–7.

33. Proshkin S, Rahmouni AR, Mironov A, Nudler E (2010) Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328:504–508.

34. Dworkin M (1996) Recent advances in the social and developmental biology of the myxobacteria. *Microbiol Rev* 60:70–102.

35. Ochi K, Kandala JC, Freese E (1981) Initiation of Bacillus Subtilis Sporulation by the Stringent Response to Partial Amino Acid Deprivation. *J Biol Chem* 256:6866–6875.

36. Ye J et al. (2010) The GCN2-ATF4 pathway is critical for tumour cell survival and proliferation in response to nutrient deprivation. *The EMBO Journal* 29:2082–2096.

37. Zhou Y, Goodenbour JM, Godley LA, Wickrema A, Pan T (2009) High levels of tRNA abundance and alteration of tRNA charging by bortezomib in multiple myeloma. *Biochemical and Biophysical Research Communications* 385:160–164.

38. Andersson SG, Kurland CG (1990) Codon preferences in free-living microorganisms. *Microbiol Rev* 54:198–210.

39. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I-1-I-2 regulatory elements. *Nucleic Acids Res* 25:1203–1210.

40. Yelverton E, Lindsley D, Yamauchi P, Gallant JA (1994) The function of a ribosomal frameshifting signal from human immunodeficiency virus-1 in Escherichia-coli. *Mol Microbiol* 11:303–313.

41. Gallant J et al. (2004) On the role of the starved codon and the takeoff site in ribosome bypassing in Escherichia coli. *J Mol Biol* 342:713–724.

Figure Legends

Fig. 1: Degeneracy lifting associated with amino acid limitation.

(A) A library of *29* variants of the yellow fluorescent protein gene (*yfp*) was synthesized. In this library, each variant (represented as a horizontal line) was designed to measure the effect of one specific codon on protein synthesis rate. The identity of this codon and that of its cognate amino acid is indicated to the left of each *yfp* variant, and the locations of this codon along *yfp* are represented as thick vertical bars. Other codons for the same amino acid that were identical across all *yfp* variants in each codon family are represented as thin vertical bars.

(B) Each *yfp* variant was constitutively expressed from a low-copy vector (SC101* *ori*, *2* copies / chromosome) in *E. coli* strains that were auxotrophic for one or more of seven amino acids.

(C) To induce amino acid limited growth, we adjusted the initial concentration of an amino acid in the growth medium to a level below that is required for reaching saturating cell density. A methyl-ester analog of the amino acid supported steady growth in the amino-acid limited phase. Growth and fluorescence curves for two *yfp* variants, CTA, red, and CTG, black, are shown as illustrative examples of degeneracy splitting upon limitation for the cognate amino acid, leucine.

(D) YFP synthesis rates during limitation for cognate amino acid – blue; YFP synthesis rates during amino acid-rich growth – grey. YFP synthesis rate was defined as the rate of fluorescence change divided by the cell density. Synthesis rates were normalized by the maximum value within each synonymous codon family, and separately in the amino acid-rich and amino acid-limited growth phases. Normalization factors (amino acid – rich, limited): Leu – 94, 81; Arg – 89, 113; Ser – 217, 343; Pro – 306, 49; Ile – 295, 45; Gln – 185, 83; Phe – 311, 20; (arbitrary units). Error bars show standard error over three replicate cultures.

Fig. 2: Altering the hierarchy of degeneracy splitting among synonymous codons.

The five leucine (arginine) tRNA isoacceptors were co-expressed together with each of the six leucine (arginine) *yfp* variants resulting in thirty tRNA-*yfp* combinations for leucine (arginine).

(A, B) Each square in the left (right) table corresponds to the difference in YFP synthesis rates of each *yfp* variant between the tRNA co-expressed strain and the parent strain without extra tRNA during leucine (arginine) limitation. YFP synthesis rates were defined in the same manner and normalized by the same factor as in Fig. 1D. YFP synthesis rate of the parent strain without extra

tRNA during amino acid limitation is shown on the left of each table (same data as in Fig. 1D). tRNA isoacceptor names are preceded by their unmodified anticodon sequences. Solid black-outlined squares correspond to codon–tRNA pairs that satisfy wobble-pairing rules after accounting for known post-transcriptional tRNA modifications (Table S9). Dashed black-outlined squares correspond to codon–tRNA pairs that do not satisfy known wobble-pairing rules but that show a significant increase in YFP synthesis rate upon co-expression of the tRNA isoacceptor. $^{UCG}Arg2_m$ is a non-native arginine tRNA that was created by mutating the anticodon sequence of the ^{ACG}Arg2 gene. Standard error was less than 0.05 for all squares.

(C) Histogram of differences in YFP synthesis rate of *yfp* variants upon tRNA co-expression. Amino acid limited growth: 42% median difference; Amino acid-rich growth: 9% median difference (n=60, aggregated for leucine and arginine). Change in YFP synthesis rate between each tRNA co-expressed strain and its parent strain expressing no extra tRNA was calculated as a percentage of the largest value between the two YFP synthesis rates.

Fig. 3: Degeneracy lifting for endogenous proteins.

(A) The effect of each codon on the synthesis rate, *S*, of a protein during amino acid limitation was modeled by a codon-specific weight, w_{codon}. The codon robustness index (CRI) for any protein coding sequence was defined as the product of w_{codon} values for all codons in that sequence that are cognate to the limiting amino acid.

(B) w_{codon} values for leucine and arginine codons during limitation for their cognate amino acids were estimated from protein synthesis rates of the corresponding *yfp* variants (Methods). w_{codon} values for all codons not cognate to the limiting acid were set to *1*.

(C) Ninety-two open reading frames (ORFs) from the *E. coli* genome were cloned as N-terminal fusions to YFP downstream a constitutive promoter into a low-copy vector (Inset, Methods). Robustness to leucine limitation is quantified as the ratio of protein synthesis rates between leucine-limited and leucine-rich growth phases. This measured robustness was correlated with estimated Leu CRI values for the *92* ORF-*yfp* fusions (r^2*=0.61*, squared Spearman rank-correlation, $P = 10^{-20}$). *11* ORFs had measured robustness below the lower limit of the vertical axis (Table S1), but were included in the calculation of r^2. Protein synthesis rates were normalized by the synthesis rate for the CTG variant of *yfp*. Error bars show standard error over three replicate cultures.

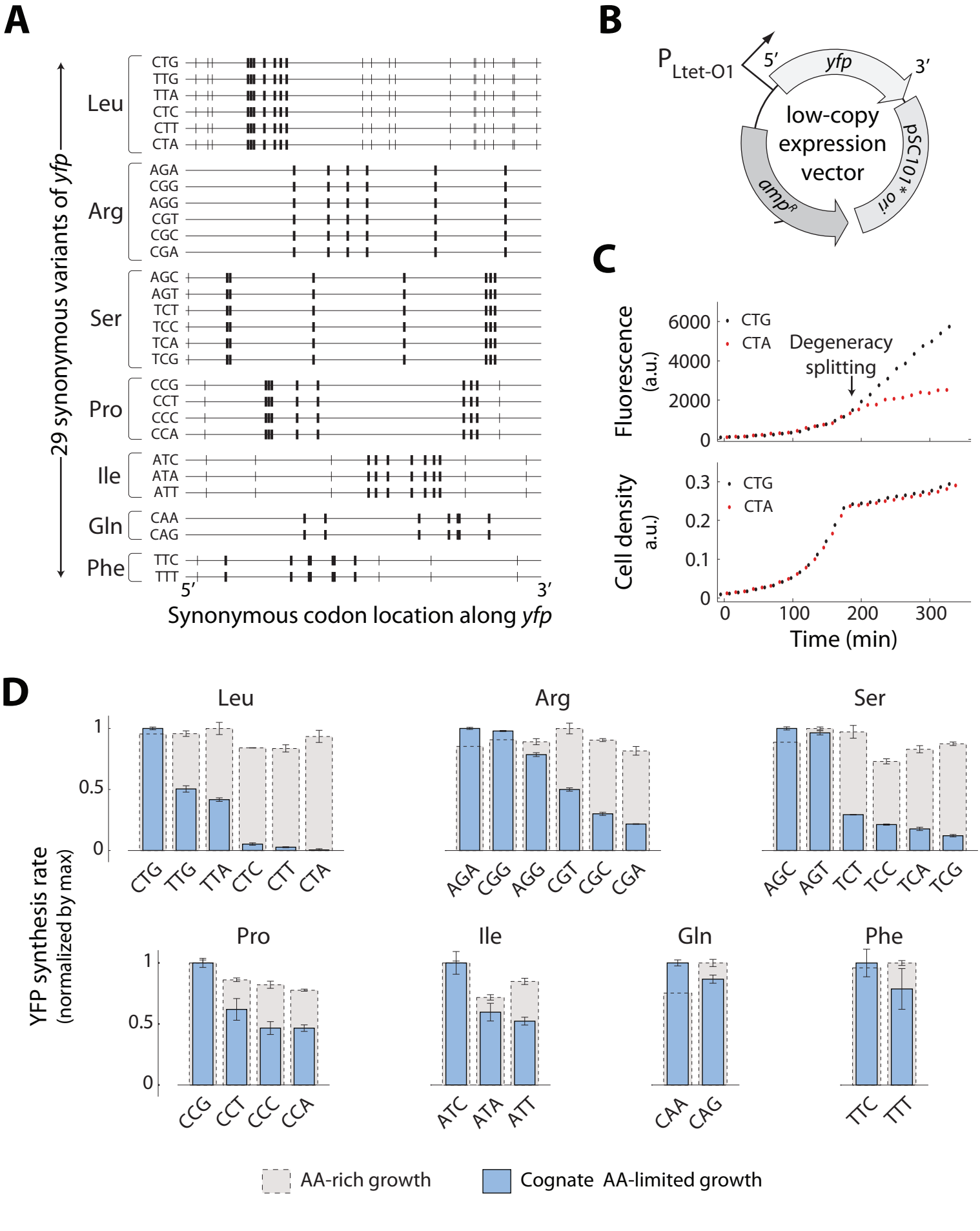
(D) Two sets of ORF-*yfp* fusions (*21* total ORFs) were co-expressed with ^{GAG}Leu2 tRNA. Based on the *yfp* data (Fig. 2A), we estimated a higher CRI for the first set (*11* ORFs) and a lower CRI for the second set (*10* ORFs) upon ^{GAG}Leu2 co-expression (Left panel, Methods). Hence we predicted that the first set should show an increase in robustness of protein synthesis during leucine limitation while the second set should show a decrease. These predictions agreed with measured changes for *20* of the *21* ORFs (Right panel, $r^2 = 0.57$, $P = 10^{-4}$). Error bars show standard error over three replicate cultures. Several error bars are smaller than data markers.

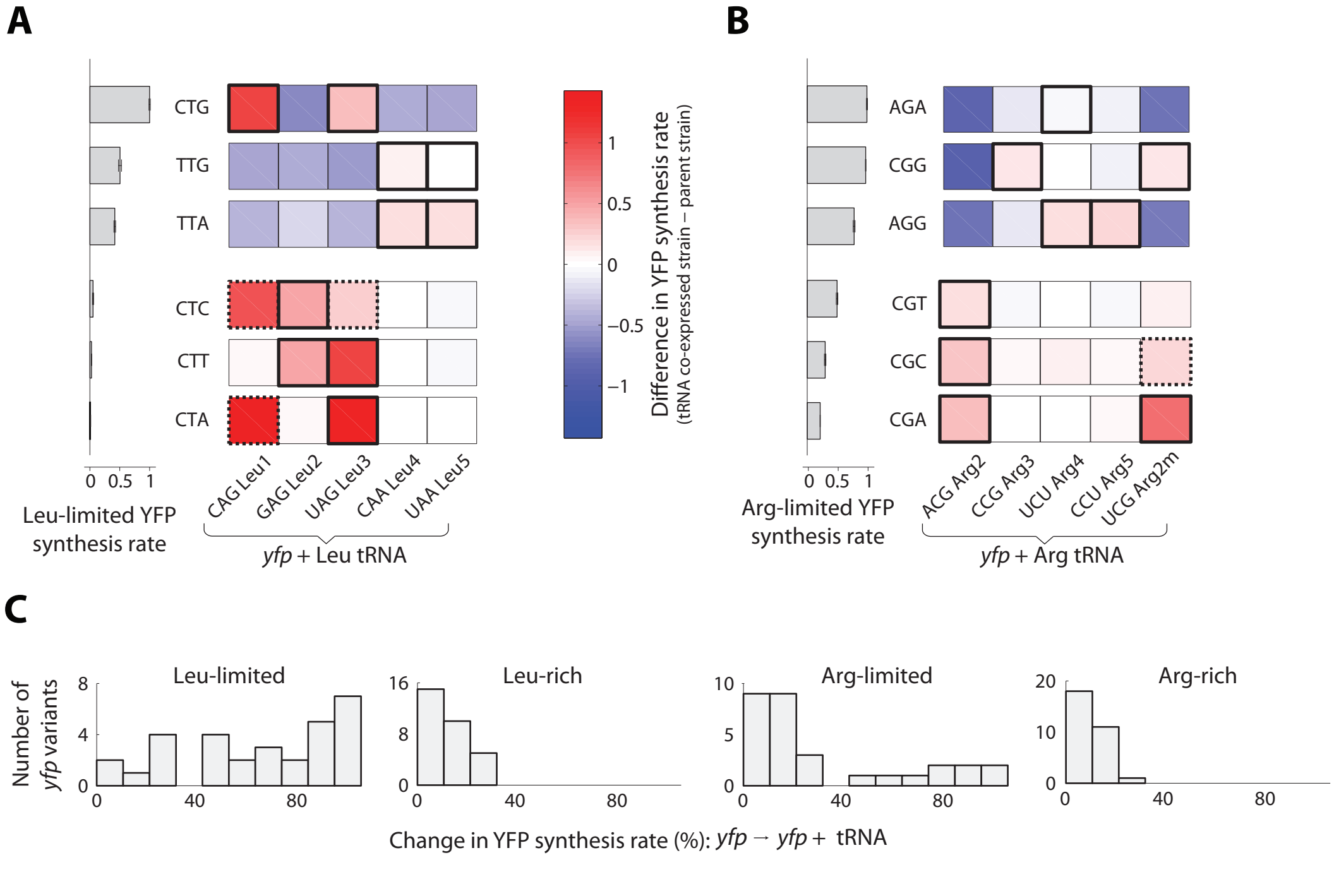
Fig. 4. Fitness cost and transcriptional control reflect degeneracy lifting.

(A) Four different prototrophic *E. coli* strains were created. Each of these strains had one of the four amino acid biosynthesis genes *argA* (Arg), *carA* (Arg), *leuA* (Leu) and *leuC* (Leu) replaced at the native locus by a corresponding synonymous mutant ORF. These mutants were designed such that three to five perturbation-robust codons in wild-type ORF were replaced by perturbation-sensitive codons in the mutant ORF (see Fig. S12, B). The strains were grown in

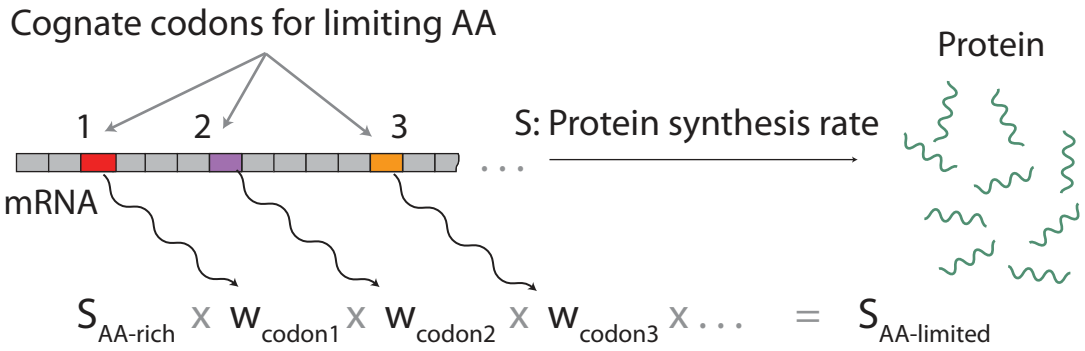
medium supplemented with all *20* amino acids at *800μM*, and then diluted into a medium lacking either leucine (*left panel*) or arginine (*right panel*). Growth lag was calculated as the time taken by each strain to reach OD_{600} of 0.3 relative to a reference culture of the same strain grown in *800μM* of all *20* amino acids. Difference in growth lag between the *leuA* mutant and the two controls during leucine downshift (left panel) was 9.2 ± 2.8 min, $P = 10^{-3}$. Difference in growth lag between the *carA* mutant and the two controls during arginine downshift (right panel) was 7.8 ± 1.2 min, $P = 10^{-6}$. Standard errors were calculated over six biological replicates for each mutant. P-values were calculated using two-tailed t-test between the *leuA* or *carA* mutant and the corresponding controls.

(B) (*Top panel*) Genes encoding sigma factors and leucine biosynthesis genes in *E. coli* are biased in their Leu CRI values, as quantified using a z-score that measures the normalized deviation from the expected CRI value based on genome-wide codon frequencies (Appendix). The most frequent leucine codon CTG was excluded in this analysis since its frequency varies significantly with expression level under nutrient-rich conditions (38). (*Bottom panel*) Fold-change in mRNA abundance in response to leucine limitation for sigma factor genes and leucine biosynthesis operons was measured using RT-qPCR. Fold-change of the *gapA* gene was used for internal normalization. Error bars show standard error over triplicate qPCR measurements.


A
29 synonymous variants of yfp
Leu
CTG
TTG
TTA
CTC
CTT
CTA
Arg
AGA
CGG
AGG
CGT
CGC
CGA
Ser
AGC
AGT
TCT
TCC
TCA
TCG
Pro
CCG
CCT
CCC
CCA
Ile
ATC
ATA
ATT
Gln
CAA
CAG
Phe
TTC
TTT
5′
3′
Synonymous codon location along yfp
B
P Ltet-O1
5′
yfp
3′
low-copy expression vector
pSC101* ori
ampR
C
Fluorescence (a.u.)
6000
4000
2000
0
CTG
CTA
Degeneracy splitting
Cell density a.u.)
0.3
0.2
0.1
0
0
100
200
300
Time (min)
D
YFP synthesis rate (normalized by max)
Leu
Arg
Ser
Pro
Ile
Gln
Phe
1
0.5
0
AA-rich growth
Cognate AA-limited growth


A
B
C
CTG
TTG
TTA
CTC
CTT
CTA
CAG Leu1
GAG Leu2
UAG Leu3
CAA Leu4
UAA Leu5
yfp + Leu tRNA
Leu-limited YFP synthesis rate
Difference in YFP synthesis rate
(tRNA co-expressed strain – parent strain)
AGA
CGG
AGG
CGT
CGC
CGA
ACG Arg2
CCG Arg3
UCU Arg4
CCU Arg5
UCG Arg2m
yfp + Arg tRNA
Arg-limited YFP synthesis rate
Number of yfp variants
Leu-limited
Leu-rich
Arg-limited
Arg-rich
Change in YFP synthesis rate (%): yfp → yfp + tRNA

A



$$w_{codon1} \times w_{codon2} \times w_{codon3} \times \ldots = \frac{S_{AA\text{-limited}}}{S_{AA\text{-rich}}}$$

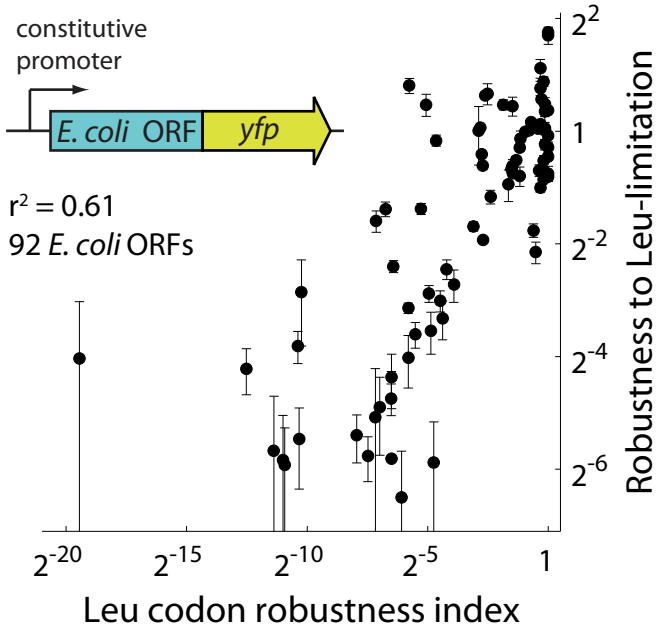
Codon robustness index (CRI) $\propto$ Robustness to AA limitation (AA-limited / AA-rich)

B

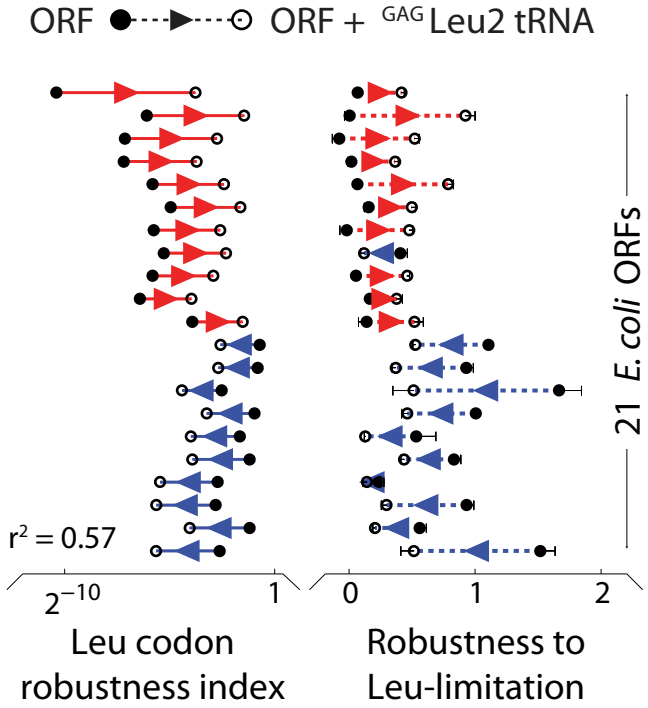
Leu	w_{codon}
CTG	1.00
TTG	0.91
TTA	0.88
CTC	0.67
CTT	0.61
CTA	0.45

Arg	w_{codon}
AGA	1.00
CGG	0.99
AGG	0.95
CGT	0.87
CGC	0.81
CGA	0.78

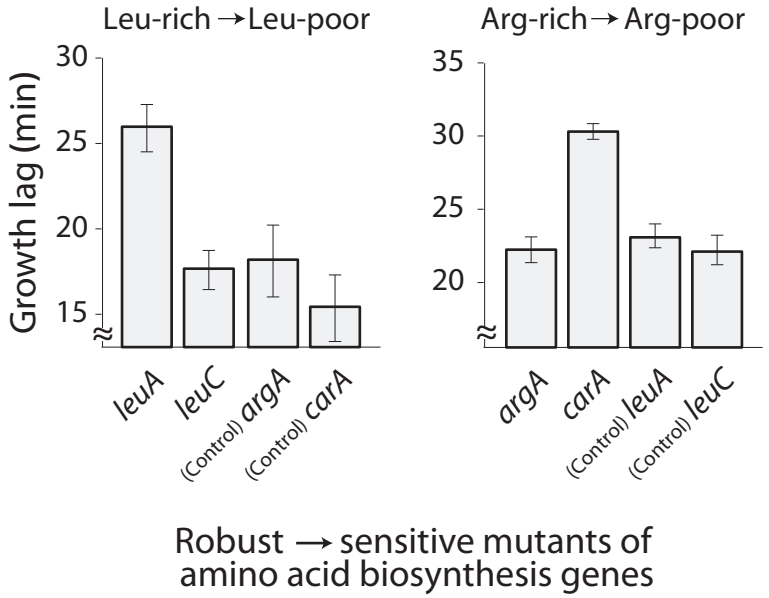
C



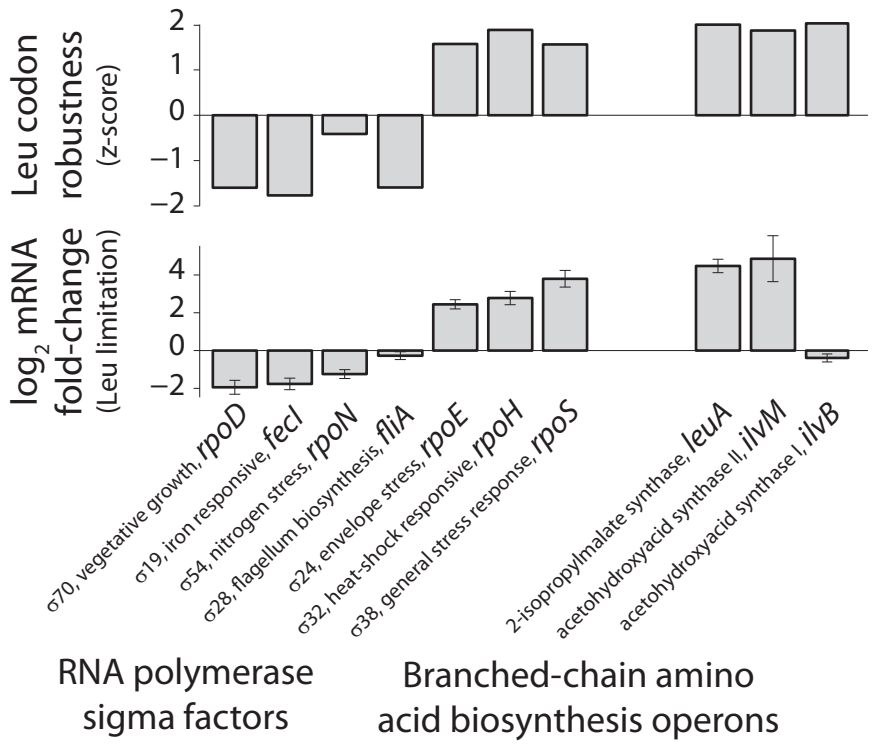
D



A


Leu-rich → Leu-poor
Arg-rich → Arg-poor
Growth lag (min)
30
25
20
15
35
30
25
20
leuA
leuC
(Control) argA
(Control) carA
argA
carA
(Control) leuA
(Control) leuC
Robust → sensitive mutants of
amino acid biosynthesis genes

B


Leu codon robustness (z-score)
2
1
0
−1
−2
log2 mRNA fold-change (Leu limitation)
4
2
0
−2
σ70, vegetative growth, rpoD
σ19, iron responsive, fecI
σ54, nitrogen stress, rpoN
σ28, flagellum biosynthesis, fliA
σ24, envelope stress, rpoE
σ32, heat-shock responsive, rpoH
σ38, general stress response, rpoS
2-isopropylmalate synthase, leuA
acetohydroxyacid synthase II, ilvM
acetohydroxyacid synthase I, ilvB
RNA polymerase sigma factors
Branched-chain amino acid biosynthesis operons

Appendix

Environmental perturbations lift the degeneracy of the genetic code to regulate protein levels in bacteria

Arvind R. Subramaniam, Tao Pan, and Philippe Cluzel

Supplementary Figures and Captions

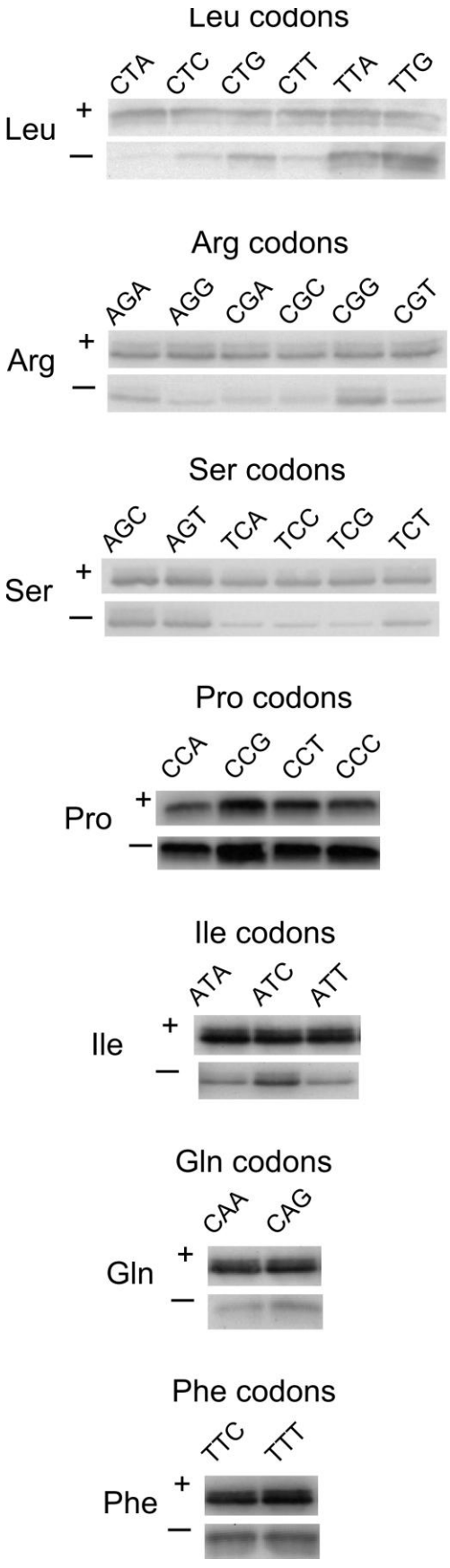

Leu codons
CTA
CTC
CTG
CTT
TTA
TTG
Leu
+
—
Arg codons
AGA
AGG
CGA
CGC
CGG
CGT
Arg
+
—
Ser codons
AGC
AGT
TCA
TCC
TCG
TCT
Ser
+
—
Pro codons
CCA
CCG
CCT
CCC
Pro
+
—
Ile codons
ATA
ATC
ATT
Ile
+
—
Gln codons
CAA
CAG
Gln
+
—
Phe codons
TTC
TTT
Phe
+
—

Fig. S1. Protein levels quantified through Western blotting reflect the difference in YFP synthesis rates during amino acid limitation.

We created modified versions of 29 *yfp* variants (Fig. 1A) that had a 3X-FLAG tag at the 5' end. These *yfp* variants were transformed into the respective *E. coli* auxotrophs in which YFP synthesis was repressed by the TetR protein (Methods). Cells were harvested at an OD_{600} of 0.4 and re-suspended in medium with or without the corresponding amino acid. Expression of YFP was induced using 200ng/ml anhydrotetracyline, and cells were harvested after 60 min. For each set of *yfp* variants under a specific growth condition, the same amount of total protein (as measured by OD_{600} before cell lysis) was used for Western blotting.

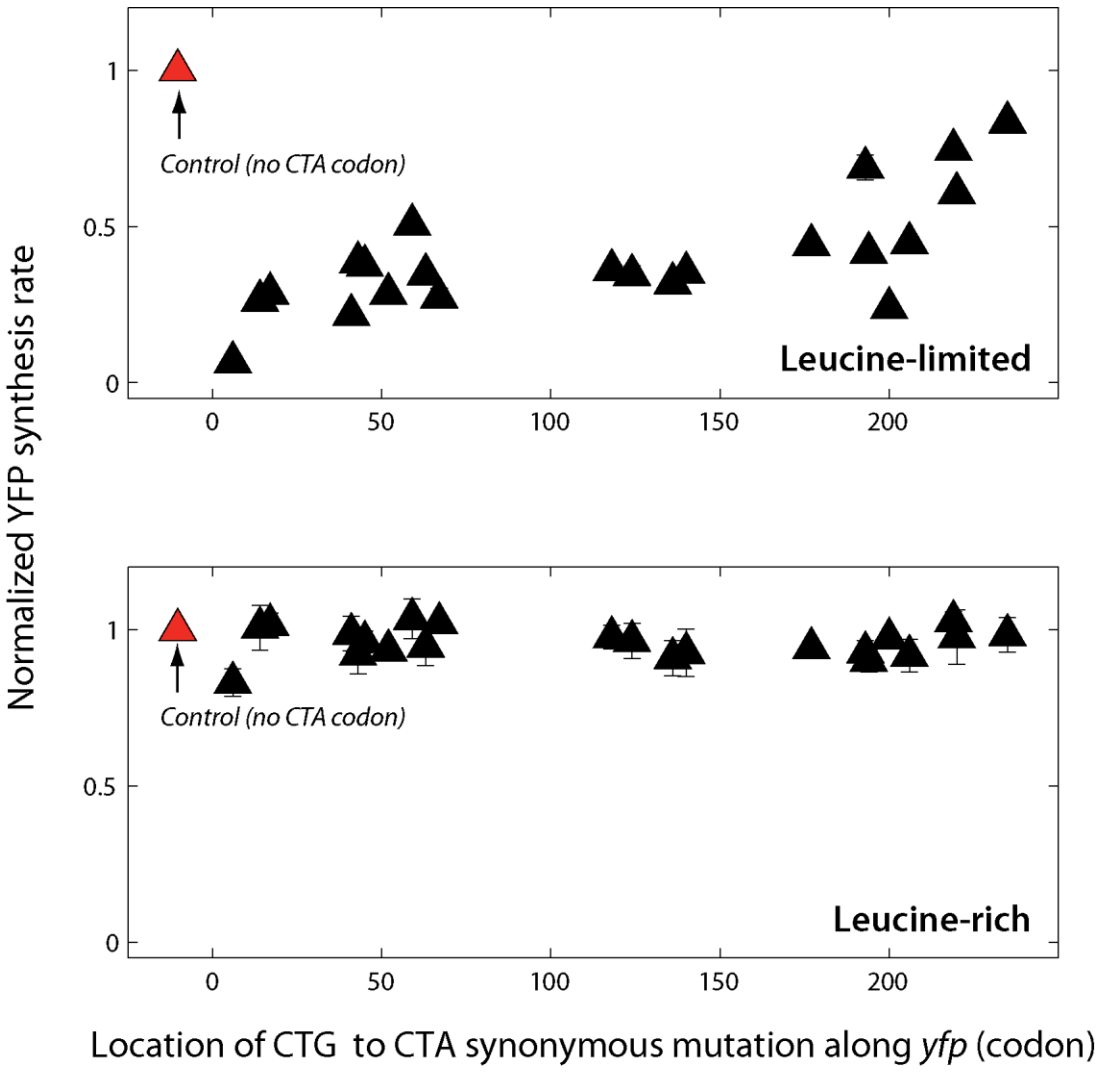

Normalized YFP synthesis rate
Control (no CTA codon)
Leucine-limited
Control (no CTA codon)
Leucine-rich
0
0.5
1
0
50
100
150
200
Location of CTG to CTA synonymous mutation along yfp (codon)

Fig. S2. Effect of a single synonymous mutation on YFP synthesis rate.

We synthesized 22 variants of the *yfp* gene each of which had a single CTA codon at one of the 22 leucine codon locations along *yfp*. The remaining leucine codons in each variant were the perturbation-robust CTG codon. The 'control' *yfp* variant did not have any CTA codon. Vertical axis refers to the YFP synthesis rate from the 22 variants normalized by that of the control variant, either during leucine limitation (top panel) or during leucine-rich growth (bottom panel). Horizontal axis indicates the location of the CTA codon along each *yfp* variant (ATG start codon = 1). Error bars show standard error over three replicate cultures.

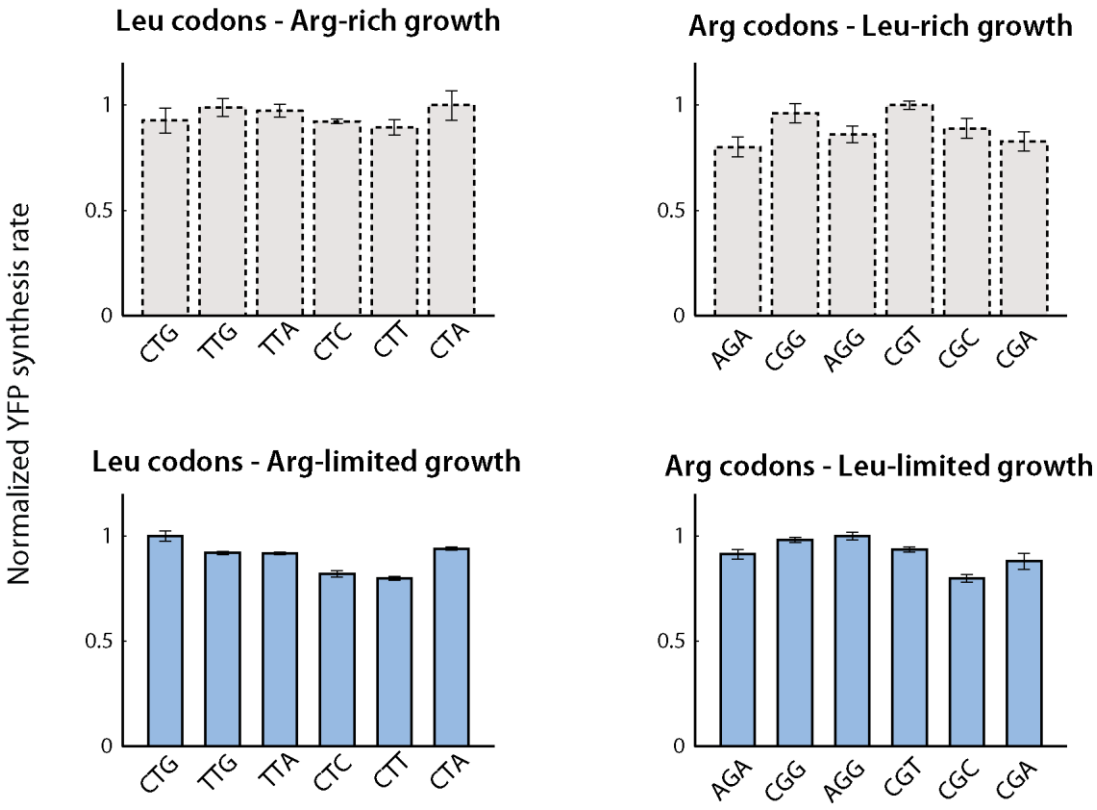

Normalized YFP synthesis rate
Leu codons - Arg-rich growth
Arg codons - Leu-rich growth
Leu codons - Arg-limited growth
Arg codons - Leu-limited growth
0
0.5
1
CTG
TTG
TTA
CTC
CTT
CTA
AGA
CGG
AGG
CGT
CGC
CGA

Fig. S3. *yfp* variants have uniform protein synthesis rates when starved for a non-cognate amino acid.

Leucine and arginine variants of *yfp* were expressed in an *E. coli* strain, CP78 that is auxotrophic for both leucine and arginine. Response of the 6 leucine variants to Arg limitation is determined by the arginine codons in *yfp* (CGT and CGC) that are common across all 6 leucine variants. Reciprocally, the response of the 6 arginine variants to leucine limitation is determined by the leucine codon that is common to the arginine variants of *yfp* (CTG). YFP synthesis rates are defined as in Fig. 1D. Error bars show standard error over three replicate cultures.

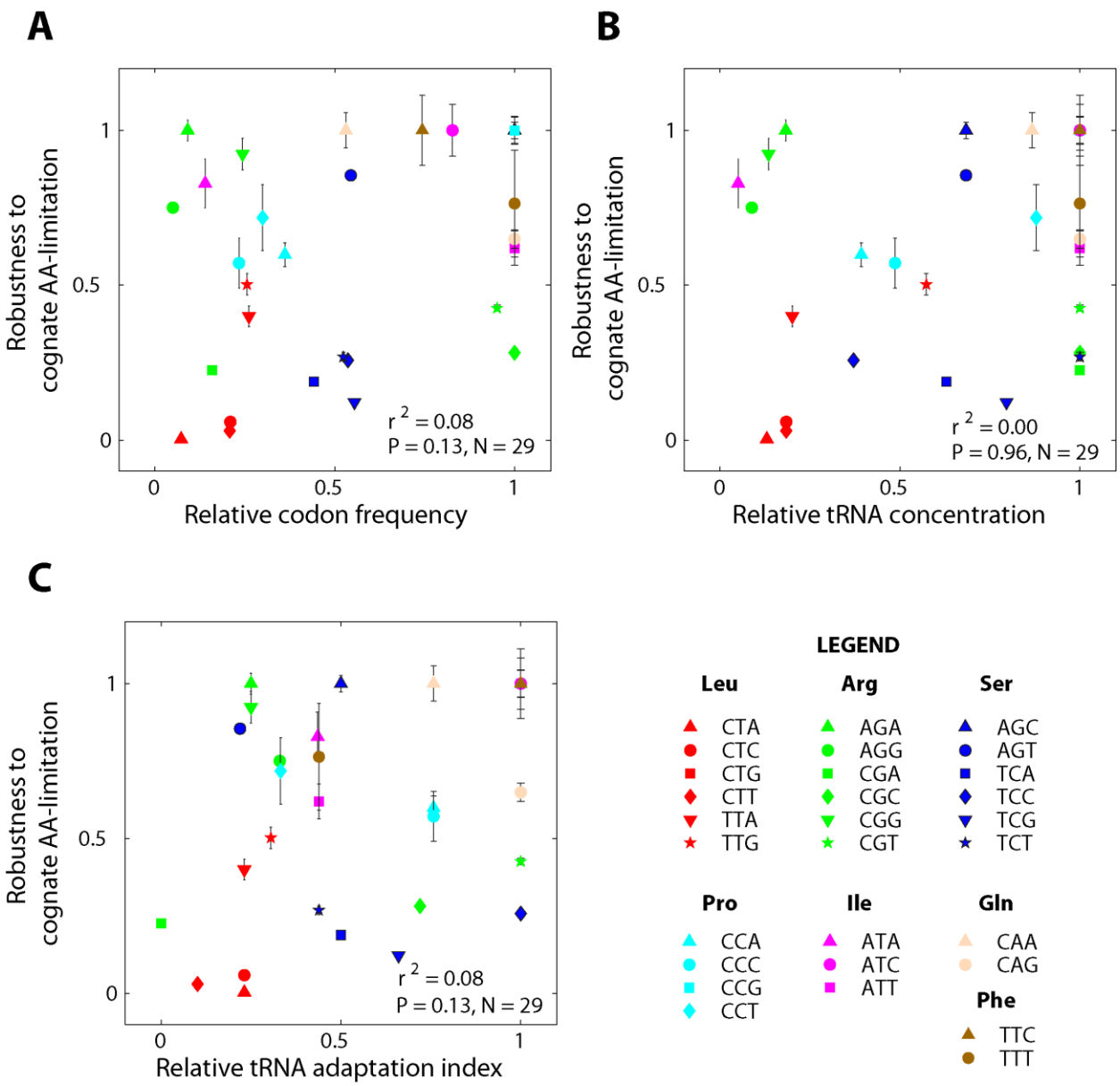

A
Robustness to cognate AA-limitation
Relative codon frequency
r 2 = 0.08
P = 0.13, N = 29
B
Robustness to cognate AA-limitation
Relative tRNA concentration
r 2 = 0.00
P = 0.96, N = 29
C
Robustness to cognate AA-limitation
Relative tRNA adaptation index
r 2 = 0.08
P = 0.13, N = 29
LEGEND
Leu
CTA
CTC
CTG
CTT
TTA
TTG
Arg
AGA
AGG
CGA
CGC
CGG
CGT
Ser
AGC
AGT
TCA
TCC
TCG
TCT
Pro
CCA
CCC
CCG
CCT
Ile
ATA
ATC
ATT
Gln
CAA
CAG
Phe
TTC
TTT

Fig. S4. Comparison of measured robustness during amino acid limitation with codon usage and tRNA concentration.

(**A**) Codon usage was calculated as the average frequency of each codon across all protein coding sequences in *E. coli*. (**B**) tRNA concentration for each codon was calculated as the sum of tRNA concentrations for all cognate tRNAs(1). (**C**) Since tRNAs can differ substantially in their affinity for their cognate codons, we also compared the measured robustness against the tRNA adaptation index for each codon (2). This index accounts for different affinities for synonymous codons for the same tRNA isoacceptor. All three measures along the horizontal axes were normalized by the maximum value within each codon family. Robustness to amino acid limitation was quantified as the ratio of normalized YFP synthesis rates between amino acid limited and amino acid rich growth phases. Error bars represent standard error over three replicate cultures. The data points that are not visible for a few codons overlap at the top right-hand corner of each plot.

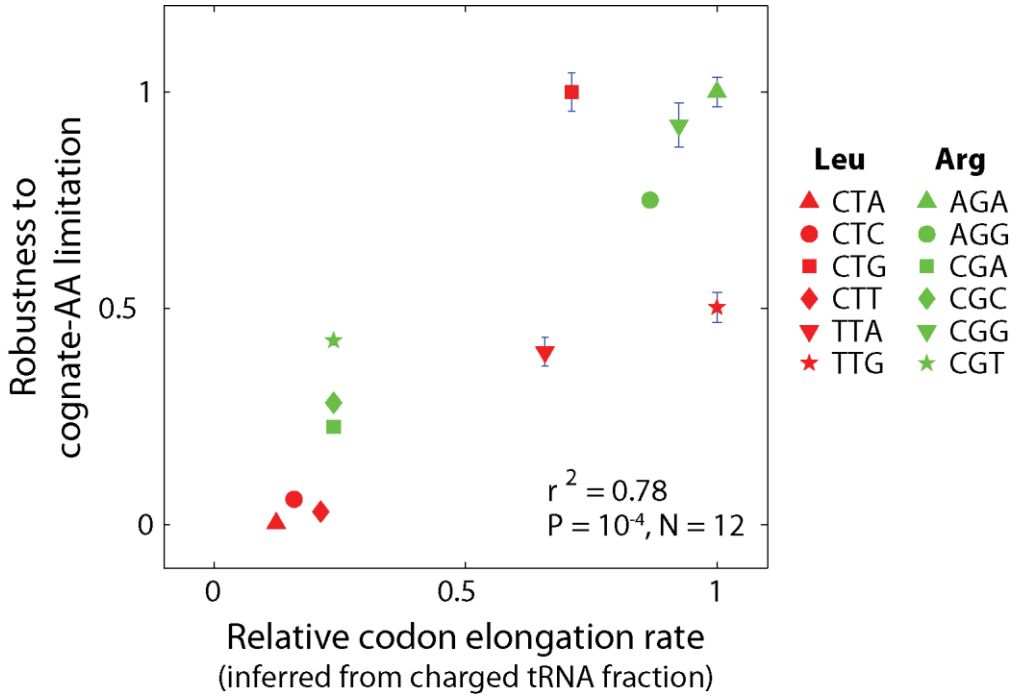

Robustness to cognate-AA limitation
1
0.5
0
0
0.5
1
Relative codon elongation rate
(inferred from charged tRNA fraction)
r 2 = 0.78
P = 10-4, N = 12
Leu
CTA
CTC
CTG
CTT
TTA
TTG
Arg
AGA
AGG
CGA
CGC
CGG
CGT

Fig. S5. Comparison of charged tRNA fraction with robustness of YFP synthesis rates during amino acid limited growth.

To compare YFP synthesis rates with charged tRNA fractions, we first inferred the elongation rates for leucine and arginine codons from the measured charged fraction of leucine and arginine tRNA isoacceptors (3) (see Appendix). We used previously assigned codon-tRNA assignments and kinetic parameters (4). Note that we cannot directly compare charged tRNA fractions with synthesis rates of *yfp* variants due to overlapping and multiple codon assignments for several tRNA isoacceptors. Robustness to amino acid limitation was quantified as the ratio of normalized YFP synthesis rates between amino acid limited and amino acid rich growth phases. Error bars represent standard error over 3 replicate cultures. Relative codon elongation rate is the ratio of codon elongation rates between amino acid starved and amino acid rich growth regimes, normalized to the maximum value within each synonymous codon family.

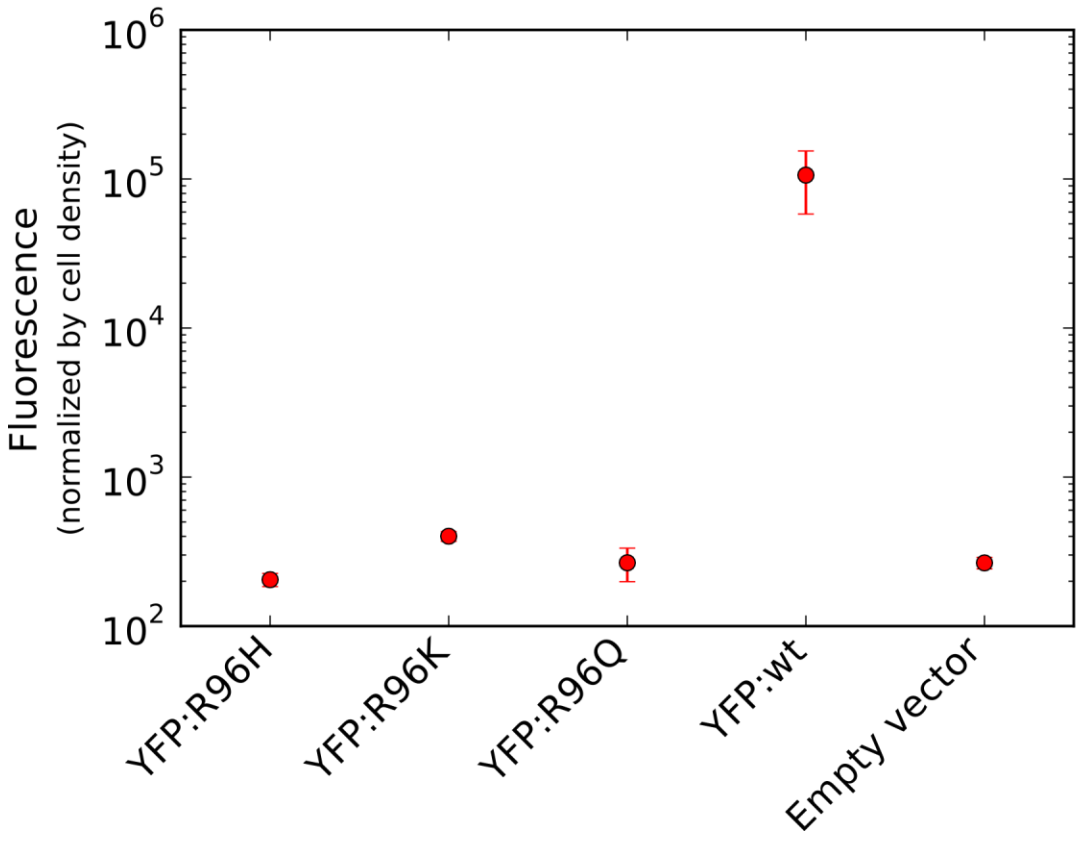

Fluorescence
(normalized by cell density)
10^6
10^5
10^4
10^3
10^2
YFP:R96H
YFP:R96K
YFP:R96Q
YFP:wt
Empty vector

Fig. S6. Miscoding of a single arginine residue in YFP causes loss of fluorescence.

To test whether mistranslation of arginine residues can underlie the high residual fluorescence of Arg *yfp* variant - Arg tRNA pairs (AGA: arg3, AGG: arg3, and CGG:arg4, arg5 in Fig. 2B), three YFP mutants were created that had one of three single point mutations at Arg 96: R96H, R96K, and R96Q. The mutant and the 'wild-type' YFP proteins were expressed from a pUC18 high-copy vector. Each of the three mutations at Arg96 to a chemically similar amino acid (H, K or Q) decreased YFP fluorescence to background level (that of an empty pUC18 vector). Error bars denote standard deviation over 5 biological replicates.

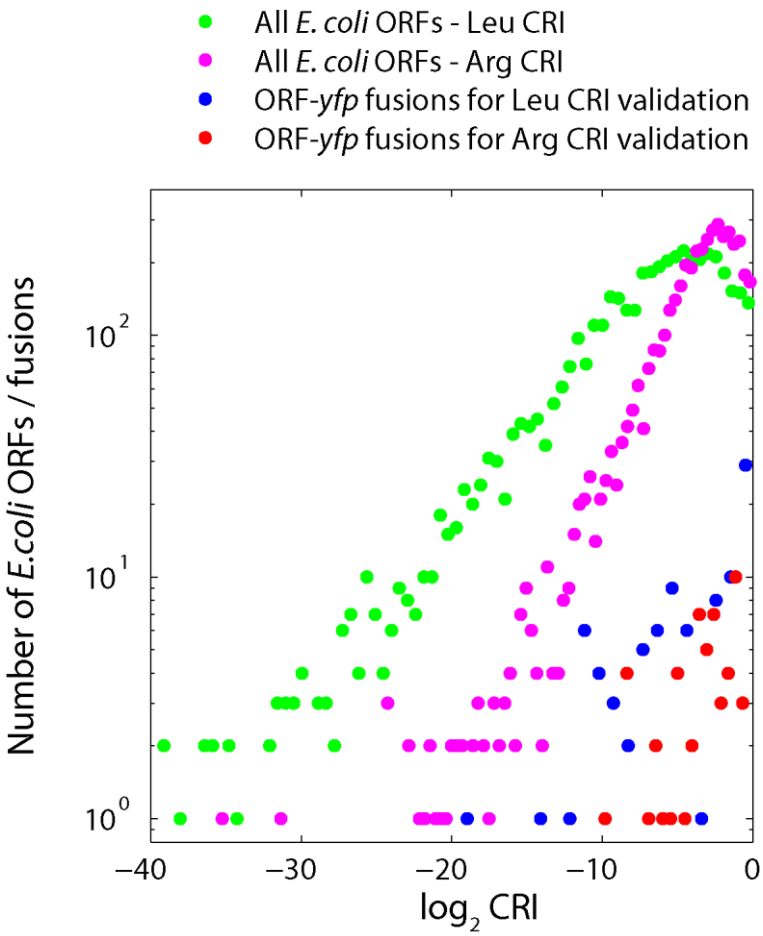

All E. coli ORFs - Leu CRI
All E. coli ORFs - Arg CRI
ORF-yfp fusions for Leu CRI validation
ORF-yfp fusions for Arg CRI validation
Number of E.coli ORFs / fusions
10^2
10^1
10^0
−40
−30
−20
−10
0
log2 CRI

Fig. S7. Histogram of CRI.

Green and pink data markers correspond to the leucine and arginine CRI values for 4300 ORFs in *E. coli*'s genome. Blue and red data markers correspond respectively to leucine and arginine CRI values for the *E. coli* ORF-*yfp* fusions respectively that were used to experimentally validate CRI.

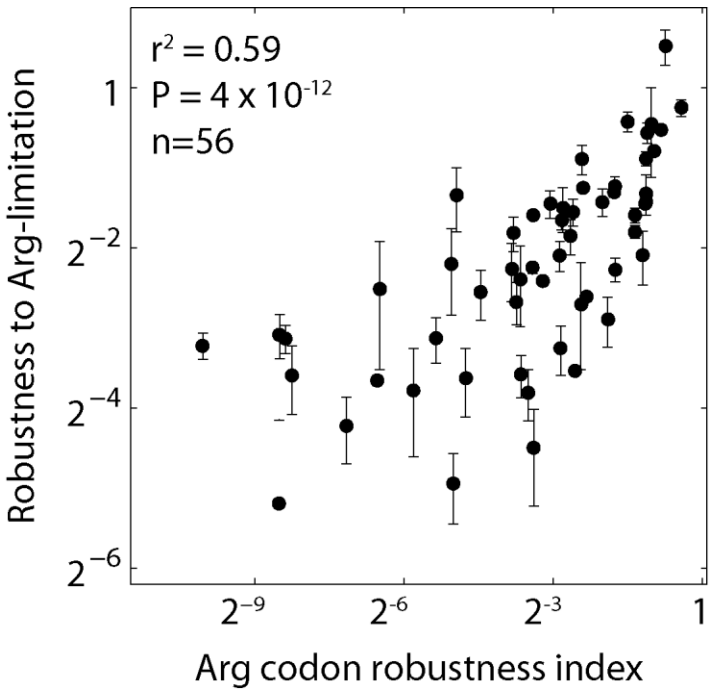

r² = 0.59
P = 4 x 10⁻¹²
n=56
Robustness to Arg-limitation
1
2⁻²
2⁻⁴
2⁻⁶
2⁻⁹
2⁻⁶
2⁻³
1
Arg codon robustness index

Fig. S8. Arg CRI is positively correlated with measured robustness of 56 E. coli ORF–yfp fusions during Arg limitation.

The *yfp* sequence used for this experiment had the AGA codon at all Arg codon locations of *yfp* since AGA has the highest w_{codon} value among arginine codons (see Fig. 3B). Correlation is reported as squared Spearman rank correlation. Error bars show standard error over three replicate cultures. . Robustness to amino acid limitation was quantified as the ratio of normalized YFP synthesis rates between amino acid limited and amino acid rich growth phases.

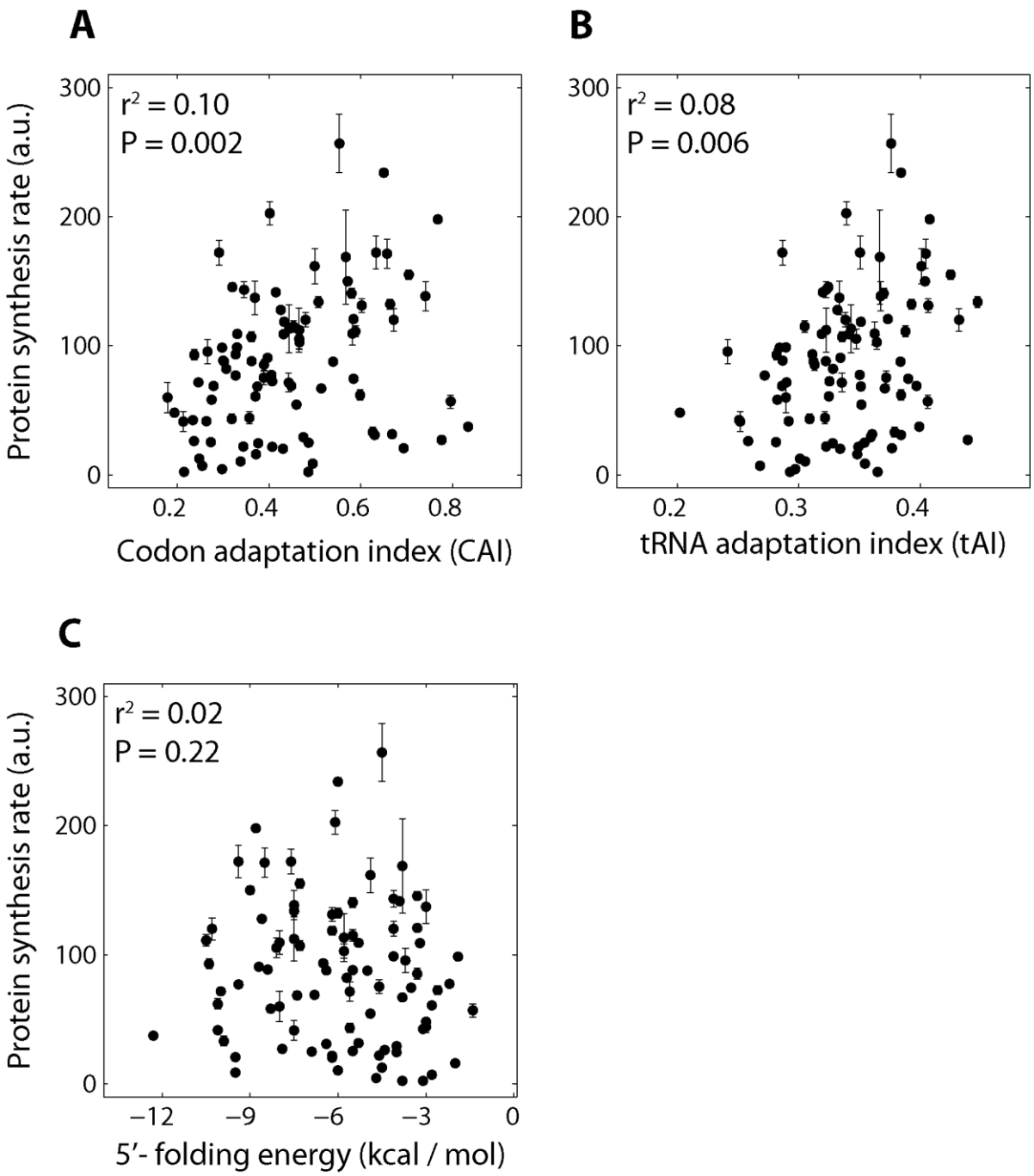

A
B
C
r² = 0.10
P = 0.002
r² = 0.08
P = 0.006
r² = 0.02
P = 0.22
Protein synthesis rate (a.u.)
Codon adaptation index (CAI)
tRNA adaptation index (tAI)
5′- folding energy (kcal / mol)
300
200
100
0
0.2
0.4
0.6
0.8
0.3
−12
−9
−6
−3

Fig. S9. Protein synthesis rate during amino acid rich growth is weakly correlated with measures of translation efficiency under nutrient-rich growth.

Protein synthesis rates from 92 *E. coli* ORF-*yfp* fusions during leucine-rich growth showed only a weak correlation with measures of codon adaptation, tRNA adaptation and 5' folding energy of mRNA (see Supplementary Methods). Folding energy was calculated from -5 to +37 nt of the ATG codon. Codon adaptation index (CAI) and tRNA adaptation index (tAI) were calculated using Biopython and codonR packages. Correlations are reported as squared Spearman rank-correlation coefficient.

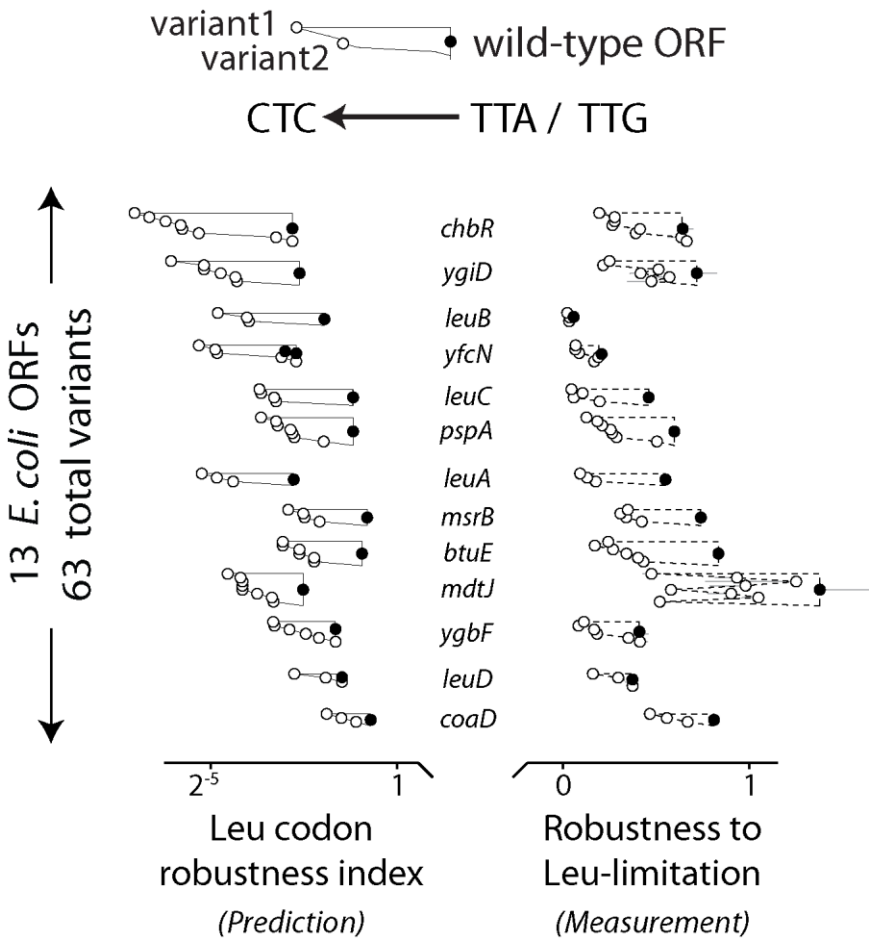

variant1
variant2
wild-type ORF
CTC
TTA / TTG
13 E. coli ORFs
63 total variants
chbR
ygiD
leuB
yfcN
leuC
pspA
leuA
msrB
btuE
mdtJ
ygbF
leuD
coaD
2-5
1
0
1
Leu codon robustness index
(Prediction)
Robustness to Leu-limitation
(Measurement)

Fig. S10. CRI predicts the change in robustness during amino acid limitation due to synonymous mutations.

63 synonymous variants of 13 ORF-*yfp* fusions were constructed by mutating wild-type TTG or TTA codons in the ORF sequence to the codon CTC that causes sensitive protein synthesis rate under leucine limitation. The number of mutations was between 1 and 6 and the location of these mutations was random. 59 of the 63 variants displayed a decrease in their robustness during leucine limitation (dashed lines) that was predicted by leucine CRI (solid lines). In addition, magnitude of the changes in robustness during leucine limitation were positively correlated with magnitude of the changes in leucine CRI ($r^2 = 0.19$, $P = 10^{-4}$). Filled circles indicate values for $ORF_{wild\text{-}type}$ and open circles indicate values for $ORF_{variants}$. Different open circles within a single polygon correspond to distinct ORF variants for the same wild-type ORF. . Robustness to amino acid limitation was quantified as the ratio of normalized YFP synthesis rates between amino acid limited and amino acid rich growth phases. Error bars show standard error over 3 replicate cultures. Most error bars are smaller than data markers. DNA sequences for variants are provided in gene_sequences.fasta supplementary file.

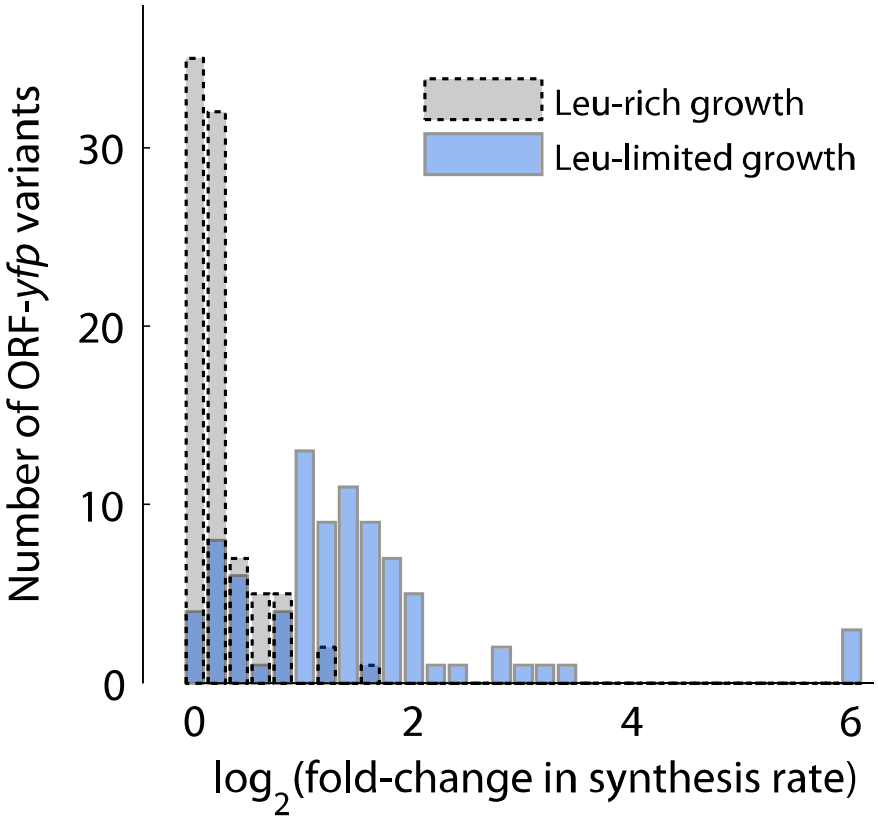

Leu-rich growth
Leu-limited growth
Number of ORF-yfp variants
30
20
10
0
0
2
4
6
log2(fold-change in synthesis rate)

Fig. S11. Synonymous mutations and tRNA co-expression affect expression from E. coli ORFs only during amino acid limitation.

We analyzed the change in protein synthesis rates from the 21 ORF-*yfp* fusions complemented with ^{GAG}Leu2 (Fig. 3D) and the 63 different ORF-*yfp* variants with synonymous mutations (Fig. S10). Several of the ^{GAG}Leu2-coexpressed as well as the synonymously-mutated ORF-*yfp* variants (84 total variants) had significantly altered protein synthesis rates compared to their non-tRNA co-expressed or non-mutated counterparts (referred as wild type) during leucine limited growth (green histogram, median fold-change in protein synthesis rates = 2.37). By comparison, most of the 84 variants had similar protein synthesis rates to their wild-type counterparts during leucine rich growth (grey histogram, median fold-change in protein synthesis rates = 1.12). Protein synthesis rates were defined as in Fig. 1D.

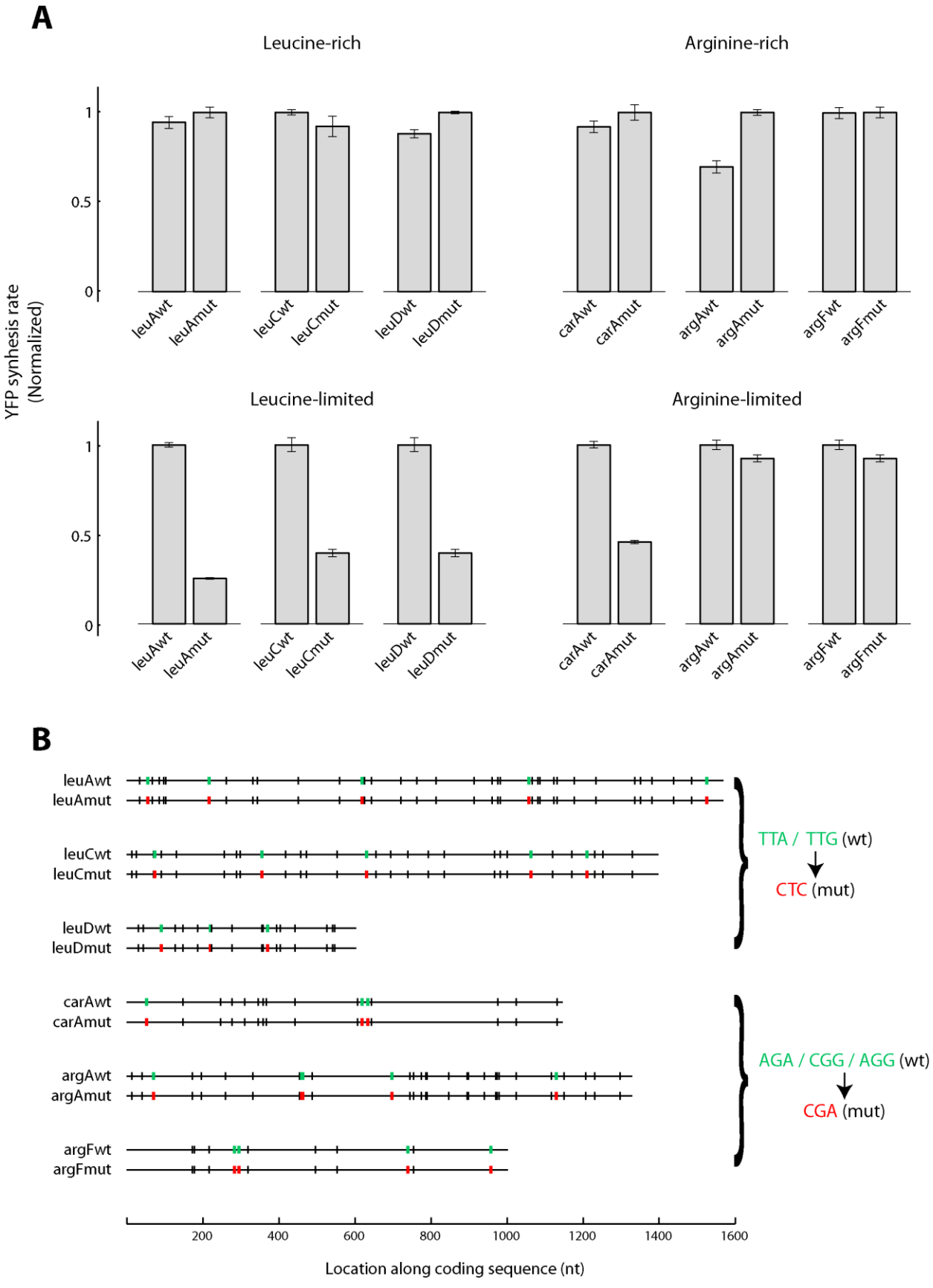

A
Leucine-rich
Arginine-rich
Leucine-limited
Arginine-limited
YFP synhesis rate
(Normalized)
1
0.5
0
leuAwt
leuAmut
leuCwt
leuCmut
leuDwt
leuDmut
carAwt
carAmut
argAwt
argAmut
argFwt
argFmut
B
TTA / TTG (wt)
CTC (mut)
AGA / CGG / AGG (wt)
CGA (mut)
200
400
600
800
1000
1200
1400
1600
Location along coding sequence (nt)

Fig. S12. Synonymous mutations in amino acid biosynthesis genes decrease their expression level during amino acid limitation.

(a) Synthesis rates from *leuA, leuC, leuD* and *carA, argA, argF -yfp* fusions encoding either wild-type or mutant ORF sequences during amino acid rich and amino acid limited growth. The synthesis rates were normalized for each pair of wild-type and mutant ORF-*yfp* fusions, and also for each growth condition. Error bars show standard error over 6 replicate cultures.

(b) Position and identity of synonymous mutations in wild-type and mutant sequences used for the experiment in (A). The black vertical bars correspond to the non-mutated leucine codons in the case of *leuA*, *leuC* and *leuD*, and to the non-mutated arginine codons in the case of *carA*, *argA* and *argF*.

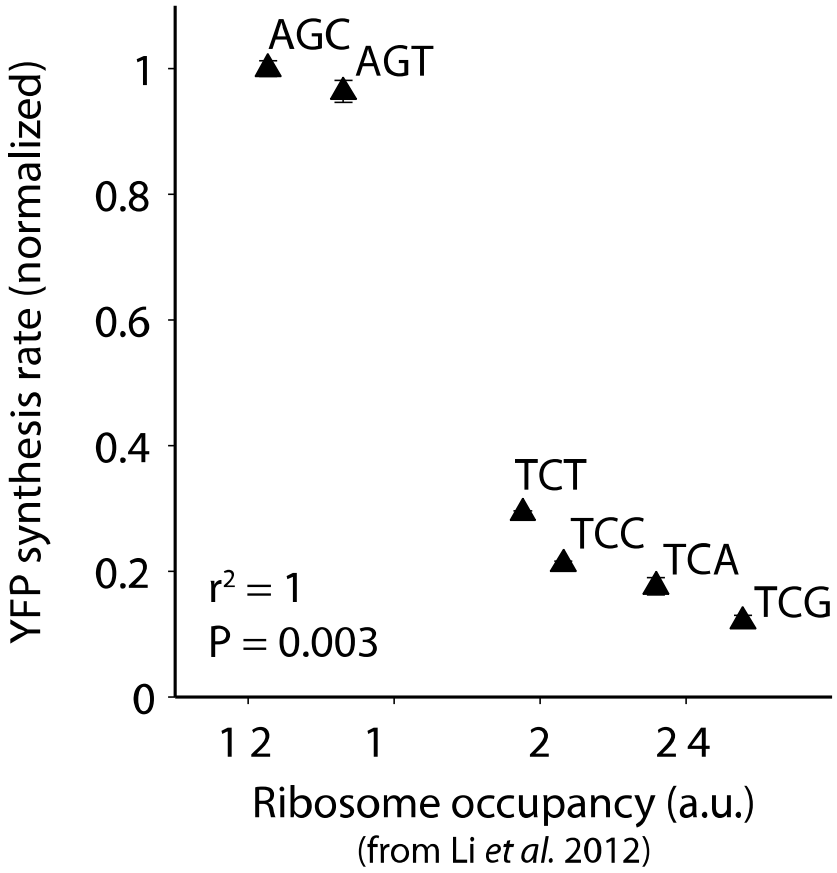

AGC
AGT
TCT
TCC
TCA
TCG
1
0.8
0.6
0.4
0.2
0
r² = 1
P = 0.003
1 2
1
2
2 4
YFP synthesis rate (normalized)
Ribosome occupancy (a.u.)
(from Li et al. 2012)

Fig. S13. Correlation of protein synthesis rates with ribosome pausing at serine codons during serine-limited growth.

Protein synthesis rate of serine synonymous variants of *yfp* during serine limitation (same data as in Fig. 1D, third panel) is negatively correlated with genome-wide ribosome occupancy at serine codons during serine-limited growth of *E. coli*. The increased occupancy at perturbation-sensitive serine codons is consistent with selective ribosome pausing at these codons. Ribosome occupancy data was taken from a recent ribosome profiling experiment in *E. coli* (5).

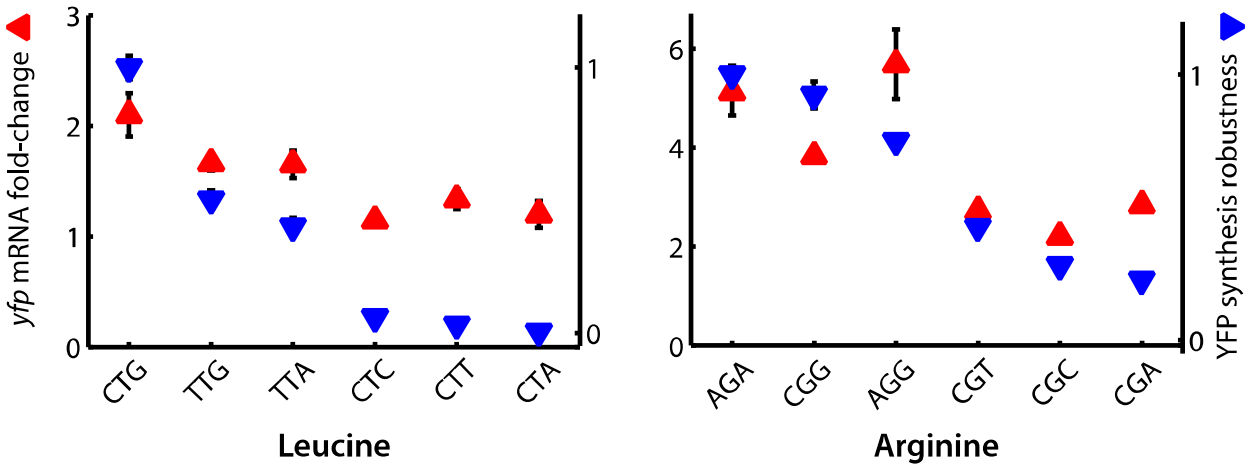

yfp mRNA fold-change
YFP synthesis robustness
3
2
1
0
1
0
CTG
TTG
TTA
CTC
CTT
CTA
Leucine
6
4
2
0
1
0
AGA
CGG
AGG
CGT
CGC
CGA
Arginine

Fig. S14. mRNA level changes of yfp variants in response to cognate amino acid limitation.

We measured the change in mRNA levels of different *yfp* variants in response to amino acid limitation. Total RNA was extracted either during exponential amino acid rich growth or 60 min after amino acid limited growth in the presence of the amino acid methyl-ester. mRNA levels were quantified by RT-qPCR relative to *gapA* mRNA (see Supplementary Methods). Error bars show standard error of triplicate qPCR measurements. . Synthesis rate robustness to amino acid limitation was quantified as the ratio of normalized YFP synthesis rates between amino acid limited and amino acid rich growth phases.

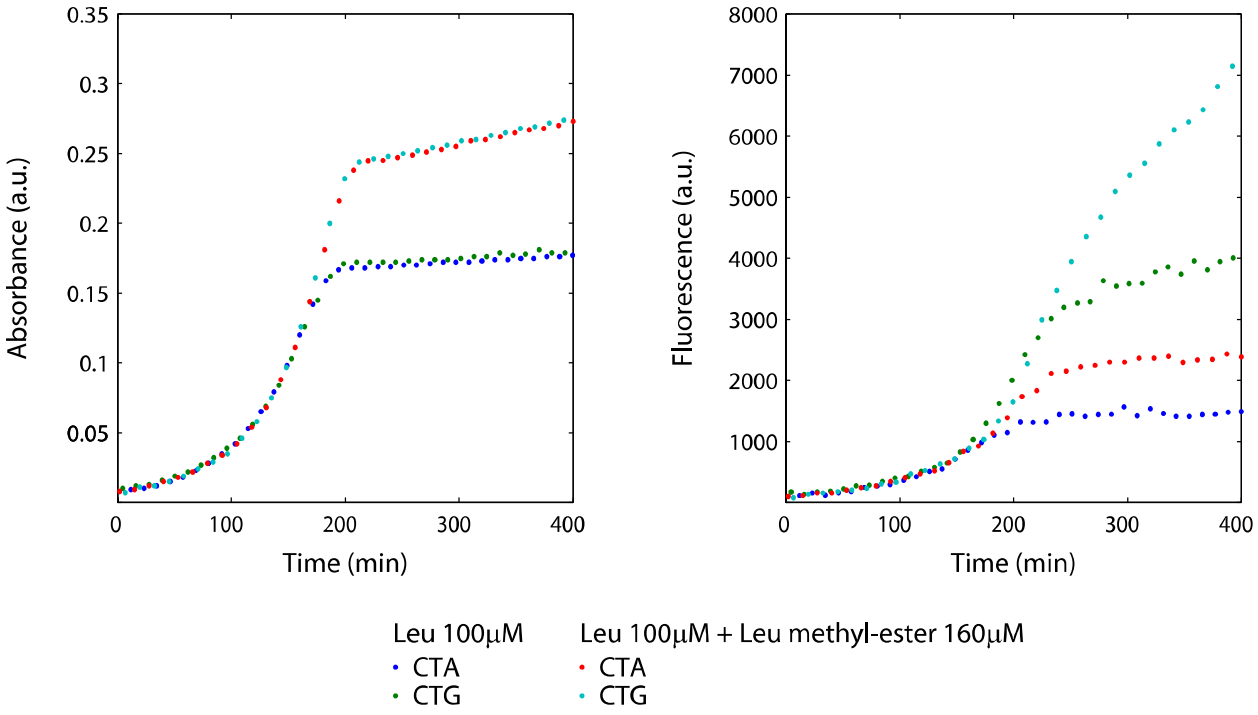
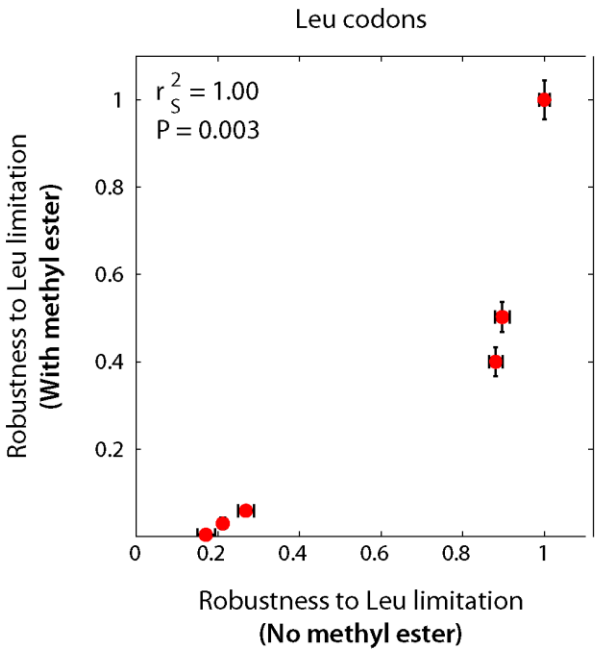

Absorbance (a.u.)
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
100
200
300
400
Time (min)
Fluorescence (a.u.)
8000
7000
6000
5000
4000
3000
2000
1000
0
100
200
300
400
Time (min)
Leu 100µM
CTA
CTG
Leu 100µM + Leu methyl-ester 160µM
CTA
CTG

Fig. S15. Raw absorbance and fluorescence curves with or without methyl-ester analog of amino acids.

Growth and fluorescence curves for two *yfp* variants corresponding to CTA and CTG codons are shown here as representative examples for amino acid limited growth in the presence or absence of methyl-ester analogs in the growth medium. Absorbance as measured using spectrometry is proportional to cell density. Presence of methyl-ester analogs caused an increase in the time and cell density at which amino acid limited growth began. More importantly, inefficient metabolism of methyl-ester analogs resulted in a slow but steady growth in amino acid limited regime. This residual growth ensured that YFP synthesis continued robustly from the CTG *yfp* variant under these conditions. By contrast, in the absence of methyl-esters in the growth medium, YFP synthesis from all *yfp* variants eventually dropped to zero.


Leu codons
$r_S^2 = 1.00$
P = 0.003
Robustness to Leu limitation
(With methyl ester)
Robustness to Leu limitation
(No methyl ester)
0
0.2
0.4
0.6
0.8
1

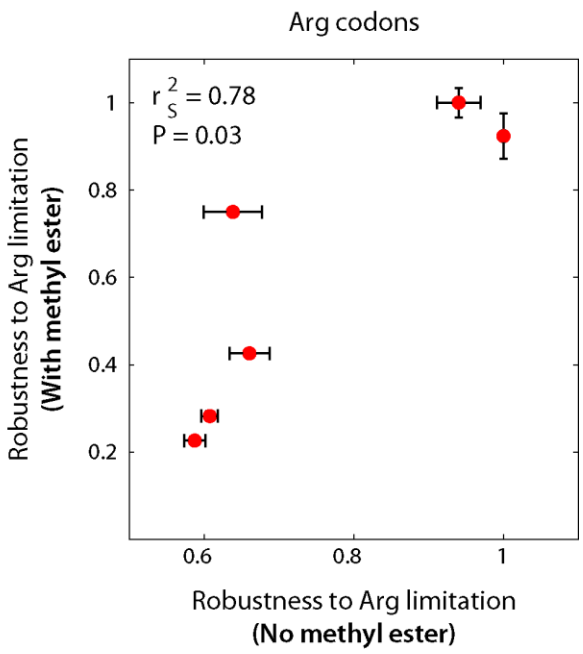

Arg codons
$r_S^2 = 0.78$
P = 0.03
Robustness to Arg limitation
(With methyl ester)
Robustness to Arg limitation
(No methyl ester)
0.2
0.4
0.6
0.8
1

Fig. S16. Comparison of two different amino acid limitation protocols: with or without methyl-ester analog of amino acids.

Robustness to amino acid limitation in the absence of methyl-ester analogs was calculated as the ratio of fluorescence change between the amino acid limited growth phase and amino acid rich growth phase. This ratio was further normalized by the maximum value within each codon family. . Robustness to amino acid limitation in the presence of methyl-ester analogs was quantified as the ratio of normalized YFP synthesis rates between amino acid limited and amino acid rich growth phases. Error bars show standard error over 3 replicate cultures.

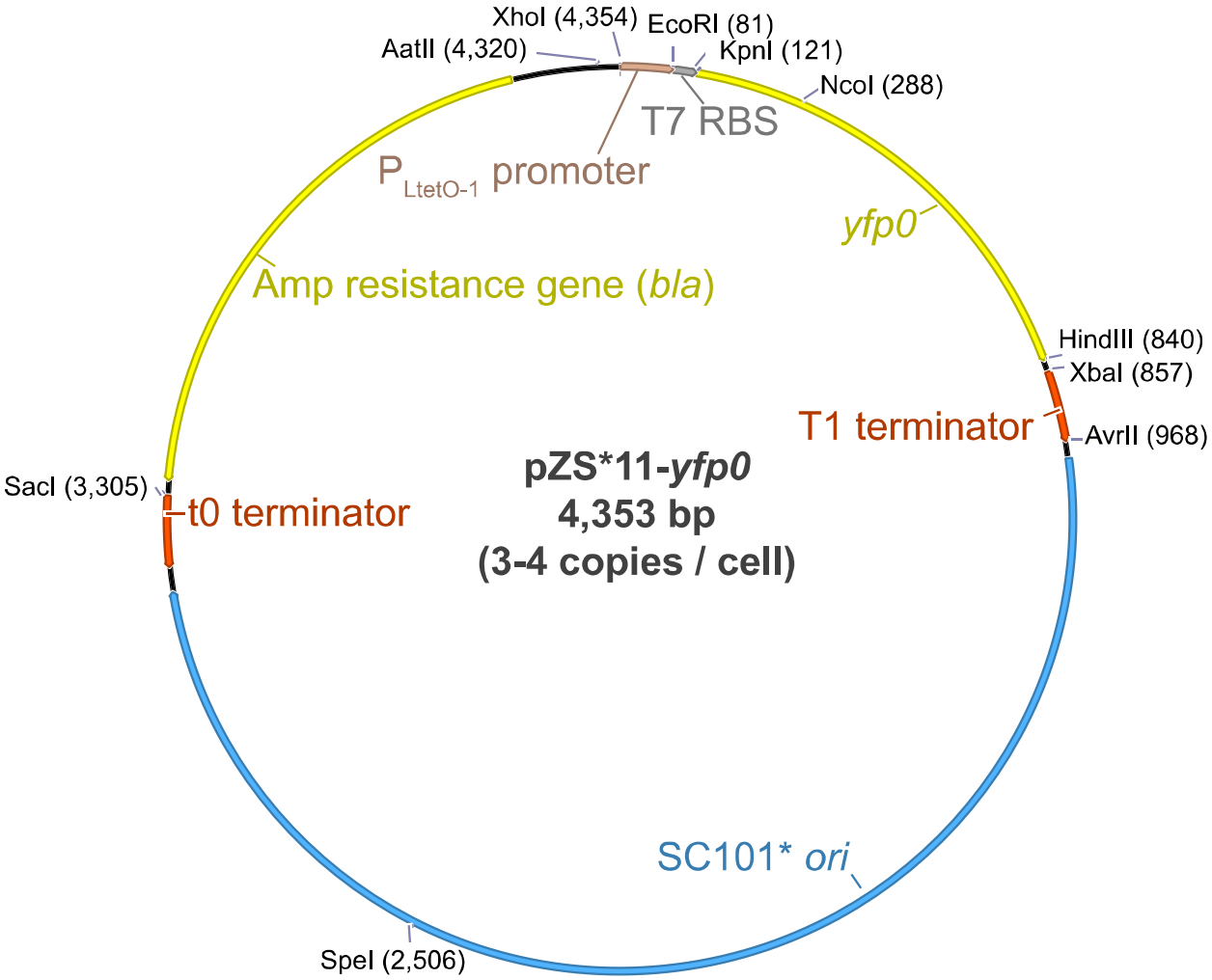

XhoI (4,354)
EcoRI (81)
AatII (4,320)
KpnI (121)
NcoI (288)
T7 RBS
P LtetO-1 promoter
yfp0
Amp resistance gene (bla)
HindIII (840)
XbaI (857)
T1 terminator
AvrII (968)
SacI (3,305)
t0 terminator
pZS*11-yfp0
4,353 bp
(3-4 copies / cell)
SC101* ori
SpeI (2,506)

Fig. S17. Plasmid map of expression vector for yfp and ORF-yfp fusions.

A specific plasmid construct with *yfp0* is shown here. In the case of ORF-yfp fusions, *yfp* was fused in-frame to the 3'-end of the ORF with a GGSGGS hexa-peptide linker sequence that encoded a BamHI restriction site and the resulting coding sequence of the fusion protein was cloned between the KpnI and HindIII restriction sites in the above vector.

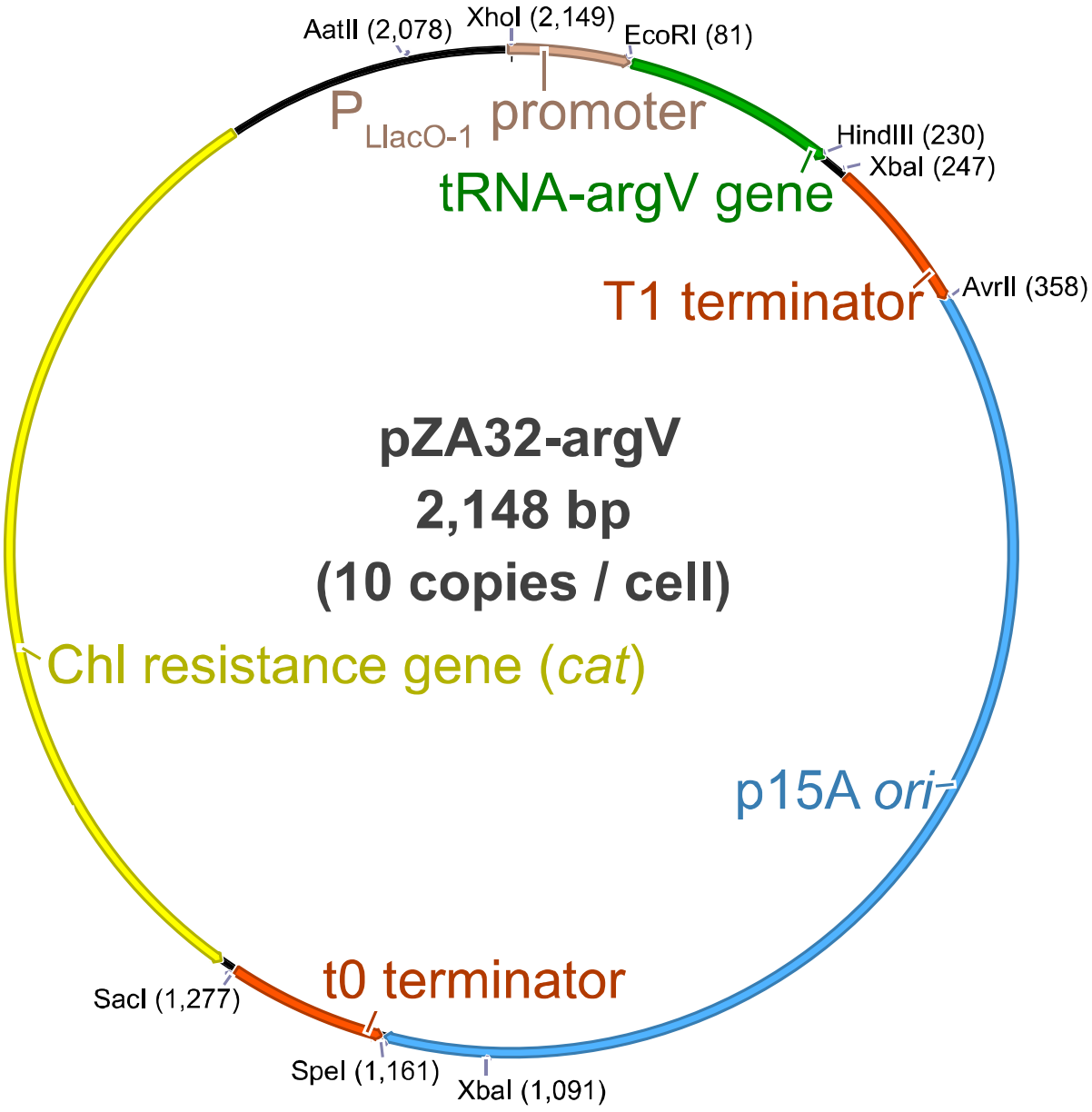

AatII (2,078)
XhoI (2,149)
EcoRI (81)
P LlacO-1
promoter
HindIII (230)
XbaI (247)
tRNA-argV gene
T1 terminator
AvrII (358)
pZA32-argV
2,148 bp
(10 copies / cell)
Chl resistance gene (cat)
p15A ori
t0 terminator
SacI (1,277)
SpeI (1,161)
XbaI (1,091)

Fig. S18. Plasmid map of expression vector for tRNA genes.

A specific construct encoding an Arg tRNA is shown here.

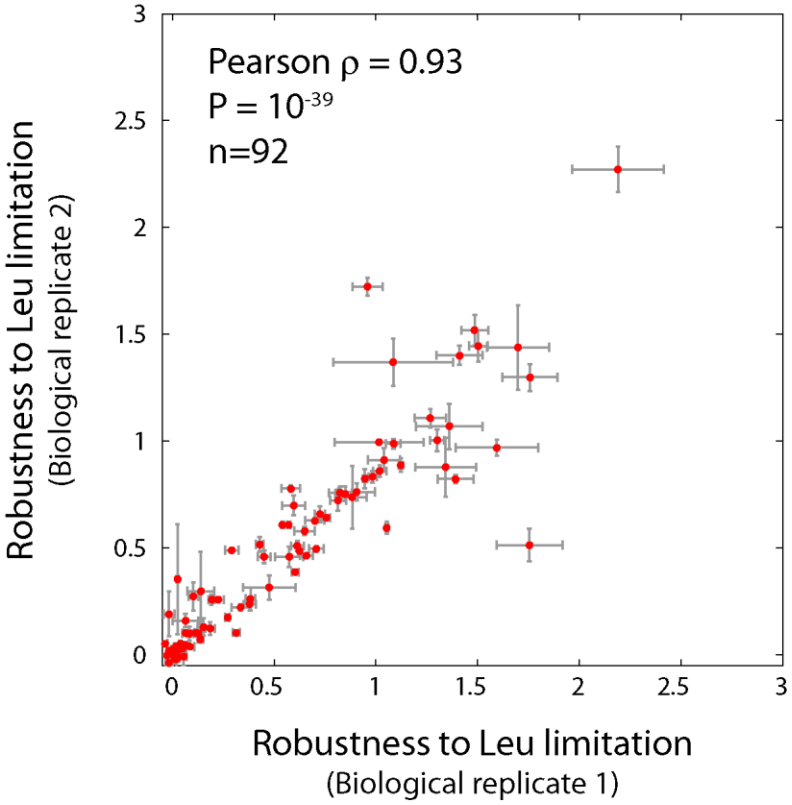

Pearson ρ = 0.93
P = 10-39
n=92
Robustness to Leu limitation
(Biological replicate 2)
Robustness to Leu limitation
(Biological replicate 1)
0
0.5
1
1.5
2
2.5
3

Fig. S19. Reproducibility of measurements between biological replicates of 92 E. coli ORF-yfp fusions.

Two different colonies were picked after cloning the 92 ORF-*yfp* fusions and the same leucine limitation assay that was used for the data in Fig. 3C was performed on these two biological replicates on two different days. None of the clones for replicate 2 were sequence-verified and hence the few outliers seen above could be the result of errors in the cloned sequences. The data reported in Fig. 3C is from replicate 1 for which about 40 constructs were sequence-verified. Robustness to leucine limitation was calculated as the ratio of normalized YFP synthesis rates between amino acid limited and amino acid rich growth phases.

Supplementary Methods

Bacterial strains

All strains used in this study were obtained from the *E.coli* Genetic Stock Center (CGSC), Yale University. For amino acid limitation experiments, standard auxotrophic strains (Table S5) were used depending on the amino acid that was limiting in the growth medium, unless mentioned otherwise. Strain CP78 was used for experiments involving leucine and arginine limitation. This strain has been used extensively in previous amino acid limitation studies (3, 6–10) and its multiple auxotrophy makes it a convenient choice for experiments involving limitation for several amino acids. The auxotrophic strains corresponding to the remaining amino acids are from the Keio-knockout collection (11), and are the commonly used auxotrophic strains for that amino acid (http://cgsc.biology.yale.edu/Auxotrophs.php).

For the growth lag measurements in Fig. 4A, the prototrophic strain MG1655 (Table S5) was used as the wild-type background. This background strain was tagged with *yfp* or *rfp* at the attBλ locus (this tagging was a remnant from earlier experiments not related to this work, and has no relevance to any results presented here). Site-directed mutagenesis was used to create the synonymous mutant coding sequences for *leuA*, *leuC*, *leuD*, *carA*, *argA* and *argF* using the protocol outlined in the section on gene synthesis and cloning below. Then to insert these mutant ORFs into their native locus without any additional markers, a two-step strategy based on λ Red-mediated homologous recombination (12) was used: In the first step, the respective wild-type ORF was replaced by a kanamycin resistance gene, and in the second step the kanamycin gene was replaced by the mutant ORF without any additional markers by selecting on M9-glucose plates for prototrophy of the respective amino acid. Plasmid pSIM5 (13) was used as the helper plasmid and the recombineering protocol outlined in the original work (13) was used without any modifications.

For the RT-qPCR data shown in Fig. 4B, a leucine auxotroph of MG1655 was created by deleting the leuB gene using the λ Red-mediated homologous recombination protocol outlined above.

For the Western blots (Fig. S1), the auxotrophic strains in Table S5 were further modified by insertion of the *tet* repressor gene at the λ-attB site using a previous method based on site-specific recombination using λ integrase (14). The presence of *tet* repressor enabled inducible control of YFP expression. The Western blots for leucine and arginine *yfp* variants were performed in an MG1655 auxotroph strain background instead of the CP78 strain. The CP78 strain has lower transformation efficiency which prevented integration of the tet repressor gene into the chromosome.

Strains were stored as 20% glycerol stocks at -80C either in 1ml cryo-vials or in 96-well plates (3799, Costar). In addition, for experiments involving over 25 strains, a temporary 20% glycerol stock was stored at -20C in 96-well PCR plates.

Plasmids

The *pZ* series of plasmids(14) were used for final expression of all genes in this study. General features of the plasmid backbones are described here. Specific gene constructs that were cloned into these backbones is described in the section on gene synthesis and cloning. A low-copy plasmid, *pZS*11* [*SC101* ori* (3-4 copies/cell), Amp^R (*bla* gene) and a constitutive $P_{LtetO-1}$ promoter] was used for expression of all fluorescent reporter genes and their fusions. The synthetic ribosome binding site (RBS) in the original *pZS*11* backbone was replaced by a modified *T7*-based RBS that resulted in efficient protein expression from most coding sequences. A medium-copy plasmid, *pZA32* [*p15A ori* (10-12 copies/cell), Chl^R (*cat* gene) and $P_{LlacO-1}$ promoter] was used for expression of all tRNA genes. Strains with *pZA32* plasmids were grown with 1mM IPTG to ensure constitutive expression of all tRNA genes. Standard plasmids *pUC18* and *pUC19* (Invitrogen) were used as intermediate cloning vectors for site-directed mutagenesis. Plasmid pSIM5 (13) was used as the helper plasmid expressing the Red system for all chromosomal modifications in this project (except for *tet* repressor insertion mentioned in the previous section on bacterial strains).

Growth and fluorescence measurements

Overnight cultures were inoculated either from freshly grown single colonies or, in experiments involving more than 25 strains, from temporary glycerol stocks stored at -20C. Overnight cultures were grown in a modified MOPS rich-defined medium (15) made with the following recipe: 10X MOPS rich buffer, 10X ACGU nucleobase stock and 100X 0.132M K_2HPO_4 (Teknova, Cat. No. M2105) were used at 1X final concentration as in the original recipe. In addition, the overnight growth medium contained 0.5% glucose as carbon source, 10^{-4}% thiamine and 800µM of 19 amino acids and 10mM of serine. pH was adjusted to 7.4 using 1M NaOH and appropriate selective antibiotics (50µg/ml ampicillin and/or 20µg/ml chloramphenicol) were added. Amino acids, glucose, thiamine and antibiotics were purchased from Sigma. 1ml overnight cultures were grown in 2ml deep 96-well plates (40002-014, VWR) at 30C with shaking at 1350rpm (Titramax 100 shaker) for 12 to 16 hours.

For experiments involving amino acid limitation, overnight cultures were diluted 1:1000 into 1ml of the same MOPS rich-defined medium as the overnight cultures. However the amino acid whose limitation was to be induced was added at a reduced concentration and supplemented with its methyl-ester analog (Table S6). Amino acid methyl-esters are analogs of the corresponding amino acids and have been previously used for steady growth of *E. coli* under amino acid limiting conditions(16, 17) (see Figs. S15 and S16 for the effect of methyl-ester on growth and robustness of YFP synthesis). Addition of the methyl-esters results in a steady but limiting supply of the amino acid due to slow hydrolysis of the ester (see Note S1). Concentrations of the amino acid and its methyl-ester were chosen such that the cultures consumed the limiting amino acid and entered amino acid-limited growth at an OD_{600} of 0.6-0.7 (corresponding to an OD_{600} value of 0.2-0.25 in our 96-well plate reader). Slight variations in the initial concentration of either the limiting amino acid or its methyl-ester shift the transition to a higher or lower cell density without appreciable changes in growth rate (see Note S2). Except for a single limiting amino acid, the remaining 19 amino acids were present at the overnight culture concentrations during the amino acid limitation experiments. For proline limitation, no proline was necessary in

the growth medium since proline methyl-ester supported growth at the same rate as proline until the OD_{600} reached around 0.6.

Diluted overnight cultures were grown in 2ml deep 96-well plates for 3 hours at 30C with shaking at 1350rpm (Titramax 100 shaker). After this time interval, 3 aliquots of 150µl each from each culture was pipetted into 3 wells of 3 different 96-well plates (3799, Costar). Wallac Victor2 plate reader (PerkinElmer) was used to monitor cell density (absorbance at 600nm) and YFP synthesis (fluorescence, excitation 504nm and emission 540nm). Each plate was read every 15 min using a robotic system (Caliper Life Sciences) and shaken in between readings (Variomag Teleshake shaker) for a total period of 6-10 hours. Temperature of 30C and relative humidity of 60% was maintained throughout the experiment.

In the case of experiments without methyl-ester (Figs. S15 and S16), the exact same protocol mentioned above was followed but the methyl-esters were not added to the growth medium.

For the RT-qPCR measurements shown in Fig. 4B, overnights cultures were diluted 1:1000 into the same medium. Then when the OD_{600} reached 0.5, the cells were spun down at 3000g for 5 min and then re-suspended in the same medium but either with or without leucine. Total RNA was extracted (see protocol below) after 30 min of shaking at 30C, 200rpm.

For the growth lag measurements shown in Fig. 4A, overnight cultures of prototrophic strains were diluted 1:200 into medium either with or without one of leucine and arginine. Growth lag was measured as the difference in time taken to reach OD_{600} of 0.3 between two cultures of the same strain – one growing in the presence of an amino acid and another growing in its absence.

Gene synthesis and cloning

All gene sequences constructed for this study are provided in the gene_sequences.fasta file. The plasmid backbone sequences are provided in the plasmid_sequences.genbank file. Primer sequences used for cloning will be provided upon request but in almost all cases, 18 to 22bp homologies without any special primer design criteria were sufficient for successful PCR amplification.

Initial yfp construct

All *yfp* variants used in this study were modified starting from a single *yellow fluorescent protein* gene sequence (called *yfp0* in the sequence file and plasmid map). This *yfp0* sequence encodes the fast-maturing 'Venus' variant of YFP (18). All 238 codons of *yfp0* were chosen such that they were read by abundant tRNA isoacceptors for each amino acid. Such a choice of codons ensured that the native level of demand for each tRNA isoacceptor inside the cell was minimally perturbed by the low-copy expression of fluorescent reporter genes. The *yfp0* sequence was built *de novo* (synthesis by Genscript, USA). The synthesized *yfp0* sequence was cloned between the *KpnI* and *HindIII* restriction sites of the *pZS*11* plasmid vector using standard molecular-biology techniques (19). The plasmid map of the resulting construct, *pZS*11-yfp0* is shown in Fig. S17.

Synonymous variants of yfp

A subset of codons in *yfp0* corresponding to 7 amino acids (Leu, Arg, Ser, Pro, Ile, Gln, Phe) were mutated to create the initial 29 synonymous variants of *yfp* (*yfp1 – yfp29* in the gene_sequences.fasta file, sequences in the same order as shown in Fig. 1A). The 4 *yfp* variants corresponding to Pro (*yfp19-yfp22*) had all the Pro codons mutated to the most frequent CCG codon since the original *yfp0* sequence had a few CCA and CCT codons that are more sensitive to Pro limitation. Similarly, all the Phe codons in *yfp0* were mutated to the most abundant Phe codon TTT for the two Phe variants of *yfp0* (*yfp28-yfp29*). Both these groups of variants (6 total) had higher overall fluorescence during amino-acid rich conditions than the rest of the 23 variants. This higher fluorescence is likely due to changes in secondary structure near the ribosome binding region on the mRNA as a consequence of mutations near ATG. However, this change is common across all variants within the Pro and Phe synonymous codon groups and hence is not responsible for the differential response to cognate amino acid limitation measured within these synonymous codon groups.

For constructing the 29 *yfp* variants, *yfp0* from *pZS*11-yfp0* was first cloned into a *pUC19* cloning vector between the *KpnI* and *HindIII* restriction sites. A commercial site-directed mutagenesis kit (Quickchange Lightening Multi, Applied Biosystems) was used to introduce the mutations corresponding to each of the 29 variants and the manufacturer's protocol was followed. The resulting variants were verified by Sanger sequencing and then cloned into the *pZS*11* expression vector backbone between the *KpnI* and *HindIII* sites. The 22 single CTA variants of *yfp* (Fig. S2) were constructed using the same procedure as above. The 29 *yfp* variants for Western blotting (Fig. S1) were created using the same procedure as above, but with the addition of a 22 codon sequence at the 5' end that encoded a 3X-FLAG peptide recognized by a commercially available, anti-FLAG, antibody (Sigma). The 22-codon sequence is: GACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGACTACAAGGATGA CGATGACAAG.

tRNA expression vectors

The 5 distinct Leu tRNA isoacceptors encoded by the genes *leuQ*, *leuU*, *leuW*, *leuX* and *leuZ*, and the 4 distinct Arg tRNA isoacceptors encoded by the genes *argV*, *argX*, *argU* and *argW* were cloned between the *EcoRI* and *HindIII* sites of the *pZA32* expression vector (Fig. S18). These genes were amplified by PCR from the chromosome of *E. coli* strain MG1655. In addition to these native tRNA genes, a synthetic tRNA gene $arg2_m$ that was cognate to the CGA Arg codon was also created. Normally, the ^{ACG}Arg2 tRNA with ICG anti-codon reads the CGA codon inefficiently through a purine-purine wobble pairing . Expressing a synthetic tRNA with an anticodon UCG restores efficient reading of this codon and is equivalent to increasing the supply of the corresponding cognate tRNA isoacceptor. This synthetic tRNA isoacceptor was created from the *pZA32-argV* expression vector by using overlap PCR to introduce the necessary single bp mutation in the anticodon of argV. The pZA32 vectors with the tRNA genes were electroporated into strains already containing the *yfp* expression vectors.

Library of E. coli ORF-yfp fusions

92 *E. coli* Open Reading Frames (ORFs) were selected for experimental validation of the Leu codon robustness index (Leu CRI). These ORFs were chosen to span a wide range of predicted Leu CRI values and functional categories (Fig. S7 and Table S1). First, a modified *pZS*11-yfp0* vector backbone was created in which the start codon of *yfp0* was replaced by a GGSGGS hexa-peptide linker sequence: GGTGGATCCGGCGGTTCT containing a *BamHI* restriction site. Next, the 92 ORFs (without the stop codon) were amplified by PCR from the chromosome of *E. coli* strain MG1655 with 5'-*KpnI* and 3'-*BamHI* restriction site overhangs. These PCR fragments were cloned into the modified *pZS*11-yfp0* vector backbone containing the *BamHI* restriction site. 13 of the 92 ORFs had either an internal *KpnI* or an internal *BamHI* site. In these cases, a larger fragment that included adjoining sections of the *pZS*11-yfp0* vector was constructed by overlap PCR and then cloned using other restriction sites (*EcoRI* or *HindIII*). Thus the final constructs had one of the 92 *E. coli* ORFs connected through a hexapeptide linker with *yfp0*. All the cloned sequences were verified by PCR for inserts of right length and 25 of the ORF constructs were verified by Sanger sequencing. Two biological replicates of each ORF construct were compared for their synthesis robustness values as measured during the amino acid limitation assay and these values showed a high degree of correlation (Pearson $\rho = 0.93$, Fig. S19).

For validating the Arg codon robustness index (Arg CRI), 56 *E. coli* ORFs that included a subset of the above 92 ORFs were chosen (Table S2). The cloning procedure was exactly analogous to the above 92 ORFs but with one difference: the *yfp0* part of the fusion construct was replaced by a synonymous variant of *yfp0* (*yfp7*) that had the Arg codon AGA instead of the CGT and CGC codons in the *yfp0* sequence. The codon AGA has the highest w_{codon} value among the Arg codons (see Fig. 3B) and hence has a minimal effect on the measured robustness of the ORF fusions during Arg limitation.

^{GAG}*Leu2-tRNA co-expression with* E. coli ORF-*yfp* fusions

Out of the 92 *E. coli* ORF-*yfp* fusions, 21 were chosen for co-expression with the ^{GAG}Leu2 tRNA that is cognate to the codons CTC and CTT. The 21 ORFs were chosen such that 11 of them had a lower Leu CRI prediction than their wild-type counterparts while the other 10 ORFs had a higher Leu CRI prediction than their wild-type counterparts. This choice also corresponded respectively to either high frequency of the non-cognate TTA and TTG codons for ^{GAG}Leu2 or high frequency of the cognate codons CTC and CTT. The strains containing the 21 ORF fusions were each made electro-competent and then transformed with the *pZA32-leuU* plasmid that expresses ^{GAG}Leu2.

Synonymous variants of *E. coli* ORF-*yfp* fusions

Out of the 92 *E. coli* ORF-*yfp* fusions, 13 were selected for creating synonymous mutants. These 13 ORFs had a high frequency of one or both of the Leu codons, TTA or TTG and these codons were mutated to the Leu codon, CTC. All these 3 codons, TTA, TTG and CTC occur at similar frequencies on average across the genome of *E. coli*. Hence mutating between these codons will not significantly change either the codon adaptation index (CAI) or the tRNA adaptation index (tAI). The 13 ORF-*yfp* fusions were amplified by PCR from the *pZS*11* vectors between the

EcoRI and *XbaI* restriction sites (see Fig. S17). These fragments were cloned between *EcoRI* and *XbaI* sites of the pUC19 cloning vector. A commercial site-directed mutagenesis kit (Quickchange Lightening Multi, Applied Biosystems) was used to introduce the TTA, TTG → CTC mutations. A unique primer was designed for each of the TTG or TTA codons in the 13 ORFs and the primers encoded the CTC mutation. All the primers corresponding to each ORF were mixed and then used in the mutagenesis reaction. This procedure resulted in mutant coding sequences with TTA, TTG - > CTC mutations at random locations. 10 colonies for each ORF were sequenced and each unique mutant sequence was then cloned into the *pZS*11* expression vector. At the end of the procedure, a total of 63 constructs were created that each had between one and seven TTA, TTG -> CTC mutations (see gene_sequences.fasta file for exact sequences).

Total RNA extraction

Total RNA was extracted for two different experiments (Figs. 4B, S14). Phenol-chloroform extraction method was used to obtain total RNA. Briefly, 3ml of cells were quickly mixed with 5ml of ice-cold water and harvested by centrifugation at 3000g for 10min. Cell pellets were re-suspended in 500 µl of 0.3M sodium acetate- 10mM EDTA pH 4.8 buffer. The re-suspended cells were mixed with 500µl of acetate-saturated phenol-chloroform at pH 4.8, 50µl of 20% SDS and 500 µl of acid-washed glass beads (G1277, Sigma). The mixture was shaken in a vortexer for 20 min at 4C. The aqueous layer was extracted twice with acetate-saturated phenol – chloroform at pH 4.8 and once with chloroform. Total RNA was precipitated with an equal volume of isopropanol and washed with 70% ethanol –50mM sodium acetate pH 4.8 and finally re-suspended in 200µl of RNase-free water. 20µl of the total RNA was treated with DNase (EN0521, Fermentas) to remove residual DNA contamination (manufacturer's instructions were followed). The DNA-free RNA was re-suspended in 200µl of RNase-free water. Intact RNA was confirmed by observation of sharp rRNA bands in native agarose gel electrophoresis.

RT-qPCR measurement

Reverse transcription (RT) was performed using 4µl of the DNA-free RNA (100-200ng) and Maxima reverse transcription kit (K1641, Fermentas), used according to the manufacturer's instructions. Random hexamer primers were used for priming the RT reaction. At the end of the RT reaction, the 20µl reaction was diluted 100-fold and 10µl of this diluted sample was used for qPCR in the next step. qPCR was performed using Maxima SYBR-Green qPCR kit (K0221, Fermentas) and manufacturer's instructions were followed. qPCR was performed in triplicates for each RT reaction and appropriate negative RT controls were used to confirm the absence of DNA contamination. *gapA* mRNA was used as internal reference to normalize all other mRNA levels. Standard curves with 6 serial dilutions were used to optimize reaction conditions and ensure amplification efficiency of between 90-100% for the *yfp* and *gapA* amplicons. $\Delta\Delta C_t$ method was used to obtain the change in mRNA levels due to amino acid limitation. The qPCR primer sequences are given in Table S8.

Western blotting

Fresh colonies were used to inoculate overnight cultures. These overnight cultures were then diluted 1:100 into 1ml of rich-defined medium with all 20 amino acids (see section on growth and fluorescence measurements for media composition). After approximately 3.5 hours of

growth at 30C when OD_{600} was ~0.4, cells were spun down at 9000g for 1 min, and then re-suspended in 1ml of rich-defined medium without the amino acid whose limitation was to be induced. This re-suspended culture was then split into two equal aliquots. The limiting amino acid was added to one aliquot (as a rich-medium control) while the other aliquot did not have the limiting amino acid. The re-suspended medium also contained 200ng/ml of anhydro-tetracycline in order to induce the pLtet-O1 promoter that controls the 3XFLAG-*yfp* variants. After growth at 30C for 60 min, cells were spun down at 12000g, 1 min and re-suspended in 40-400ul of CellLytic B buffer (Sigma, B7435). The buffer volume used was proportional to the OD_{600} measured at the time of harvesting the culture. The lysate was stored at -80C. 10ul of the lysate was mixed with 2X Laemmli Buffer (Biorad) and then loaded onto each lane of a pre-cast polyacrylamide gel (Biorad) and SDS-PAGE was carried out at 100V for 120 min. Proteins were transferred to a nitrocellulose membrane by semi-dry blotting at 180mA for 60 min. The membrane was blocked in 2% skim-milk-TBST overnight, and then incubated with a 1:2000 dilution of an anti-FLAG antibody (F3165, Sigma) in 2% skim-milk-TBST with shaking at room temperature for 90 min. After washing 4 times for 5 min with TBST, the membrane was incubated with 1:2000 dilution of a secondary HRP-conjugated antibody (7076, Cell Signaling) in 2% skim-milk-TBST with shaking at room temperature for 60 min. After washing 4 times for 5 min with TBST, the membrane was treated with an HRP substrate (L00354, Genscript) for 5 min and exposed for 30s to a luminescence imager.

Analyses – data, models, bioinformatics

Matlab R2009b (Mathworks) was used for all analyses unless otherwise mentioned. All correlations and P-values reported in this work were calculated using the Matlab command 'corr' with the 'Type' option set to either 'Spearman' or 'Pearson' as appropriate.

Growth and fluorescence analysis

Background absorbance and fluorescence values (obtained from wells containing only growth media) were subtracted from the measured time series for each well. An exponential curve was fitted to the amino acid-rich growth regime for all data points located at least 50 min before the onset of amino acid limitation. A straight line was fitted to the amino acid-limited growth regime for all data points located at least 50 min after the onset time. These fits were performed using the Matlab command 'fit', and the in-built library options 'Exp1' and 'Poly1' respectively. To automatically identify the onset time, the intersection point between the two fitted curves was designated as the onset time of amino acid limitation. This inferred onset time coincided with the onset time identified through visual inspection of the growth curves.

To minimize noise in calculated protein synthesis rates, an exponential curve was fitted to the amino acid-rich regime of the fluorescence time-series and a straight line was fitted to the amino acid-limited regime of the fluorescence time-series. These fits were performed using the Matlab command 'fit', and the in-built library options 'Exp1' and 'Poly1' respectively. Protein synthesis rate, S was calculated as:

$$\text{Protein synthesis rate}, S = \frac{1}{\text{Absorbance}} \times \frac{d\,(\text{Fluorescence})}{d\,(\text{time})}$$

First, the above formula was evaluated at the onset time of amino acid limitation using the exponential fits for absorbance and fluorescence data in the amino acid rich growth regime. Next, the same formula was evaluated at the onset time using the linear fits for absorbance and fluorescence data in the amino acid limited growth regime. These two values correspond to the protein synthesis rates reported for the amino acid rich and amino acid limited growth regimes (such as the data in Fig. 1D). The protein synthesis rates were normalized within each synonymous codon family and for each growth condition. Robustness of protein synthesis to amino acid limitation was calculated as the ratio of normalized protein synthesis rates between the amino acid rich and amino acid limited growth regimes.

In the case of the experiment without methyl-ester (Fig. S15 and Fig. S16), the onset time of amino acid limited growth was determined exactly as above. Then starvation robustness was calculated as the normalized ratio of total fluorescence increase after the onset of amino acid limited growth to the fluorescence increase before this onset. Total fluorescence increase rather than protein synthesis rate was used for this analysis since protein synthesis rates without methyl-ester decreased continuously to zero after the onset of amino acid limited growth.

Calculation of CRI (Fig. 3)

CRI for a protein coding sequence corresponding to a limiting amino acid was calculated by multiplying the w_i weights for codons cognate to the limiting amino acid in that sequence. w_i values that are shown in Fig. 3B were calculated using the robustness of protein synthesis of the corresponding *yfp* variants during cognate amino acid limitation (Fig. 1D). Based on our non-cognate limitation experiment (Fig. S3), the w_i values for all codons other than those cognate to the limiting amino acid are set to be equal to 1.

For illustration, we demonstrate the calculation of w_i for the 6 leucine codons below. The exact same procedure was followed for other synonymous codon families. Taking $\log_2 w_i \equiv W_i$ for each codon, and $\log_2$ robustness to amino acid limited growth $\equiv SR$ for each *yfp* variant,

$$7 \times W_{CTA} + 15 \times W_{CTG} = SR_{yfp,CTA}$$

$$7 \times W_{CTC} + 15 \times W_{CTG} = SR_{yfp,CTC}$$

$$22 \times W_{CTG} = SR_{yfp,CTG}$$

$$7 \times W_{CTT} + 15 \times W_{CTG} = SR_{yfp,CTT}$$

$$7 \times W_{TTA} + 15 \times W_{CTG} = SR_{yfp,TTA}$$

$$7 \times W_{TTG} + 15 \times W_{CTG} = SR_{yfp,TTG}$$

The multiplicative factors on the LHS in front of W_i correspond to the number of different leucine codons in the corresponding leucine variant of *yfp* (see Fig. 1A, left panel). The RHS is the measured ($\log_2$) robustness of protein synthesis from the corresponding *yfp* variant during leucine limitation (see Fig. 1D, top left panel). These equations were solved simultaneously to determine the w_i value for each leucine codon. Revised w_i values (Table S7) based on *yfp*

measurements in the presence of ^{GAG}Leu2 tRNA were used for calculation of leucine CRI in the case of ^{GAG}Leu2 tRNA co-expression with *E. coli* ORFs (Fig. 3D).

Leu and Arg CRI for E. coli ORFs (Fig. 3)

4300 ORF sequences for *E. coli* were parsed out from the MG1655 genome sequence (NCBI website, Accession number: NC_000913, downloaded on 14th Apr 2011). For each of these 4300 *E. coli* ORFs, Leu or Arg CRI was calculated by multiplying the w_i values for either all Leu or all Arg codons respectively in the ORF sequence. For the 63 synonymous variants of 13 ORFs (Fig. S10), Leu CRI values were calculated using the same procedure as above after accounting for the synonymous mutations. For the 21 ORFs co-expressed with Leu2 tRNA (Fig. 3D), revised w_i values were first calculated using the method outlined in the previous section (Table S7), and using measurements on the 6 Leu variants of *yfp* complemented with ^{GAG}Leu2 tRNA(3rd column in Fig. 2A). These revised w_i values were then used to calculate Leu CRI under tRNA co-expression for the 21 tRNA co-expressed ORFs applying the same procedure as for the non-co expressed case.

Z-score for codon robustness index (Fig. 4B)

To quantify the deviation in codon robustness index from its expected value for each of the 4300 ORFs in the *E. coli* genome, 1000 random coding sequences were generated for each ORF. Each random version preserved the original amino acid sequence, but the codons for a single amino acid were sampled randomly from a multinomial distribution based on the average frequency of codons for that amino acid across the genome. CRI values were calculated for each random version of the gene, and a distribution of CRI values were generated from the 1000 random trials. The average, μ_{CRI} and standard deviation, σ_{CRI} of this CRI distribution was used to calculate the Z-score for CRI as follows:

$$Z_{CRI} = \frac{CRI_{observed} - \mu_{CRI}}{\sigma_{CRI}}$$

In the case of the Z-score for leucine shown in Fig. 4B, the leucine codon CTG was not randomized in the above calculation and only the remaining 5 leucine codons: CTA, CTC, CTT, TTA, and TTG were randomized based on their genome-wide codon frequencies. This step is essential since CTG, which is read by an abundant tRNA isoacceptor, is enriched in highly-expressed genes and such genes will show up falsely as perturbation-robust genes because CTG is also the codon that is most robust to leucine limitation in our experiments (see Fig. 1D).

Codon-specific bioinformatic measures (Figs. S4 *and* S5)

Codon usage (Fig. S4A) was calculated as the average frequency of each codon across the genome of *E. coli* MG1655 (4300 ORFs).

tRNA concentrations (Fig. S4B) were taken from previous work(1) (see Table 2). Concentrations of all cognate tRNAs for each codon were summed together.

The codon-tRNA adaptation index (Fig. S4C) is taken from literature ((20), see Table S2). The tAI value for the CGA codon was revised from the unrealistically low value of 0.00005 to 0.1333 as explained previously (21).

For inferring codon elongation rates from charged tRNA fractions (Fig. S5), we used the formula for codon elongation rate from (4):

$$\frac{1}{v_k} = \tau_0 + \frac{1}{t_i \alpha_i k_{ik}},$$

where v_k is the elongation rate of codon k, τ_0 is the codon-independent elongation time across any codon, t_i is the concentration of tRNA isoacceptor i that is cognate to codon k, k_{ik} is the second-order rate constant for binding of the ternary complex containing the charged isoacceptor i to the ribosome at codon k, and α_i is the charged fraction of isoacceptor i. We calculated the codon elongation rates during amino acid limitation using the measured charged fractions from (3). For amino acid rich conditions, we set the charged fraction to be equal to unity. We used $\tau_0 = 0.05\ s^{-1}$, and $k_{ik} = 2\ \times 10^7 M^{-1} s^{-1}$ similar to (4). The ratio of codon elongation rates was then normalized within each codon family by the maximum value within that family.

ORF-specific bioinformatic measures (Fig. S9)

The measured protein synthesis rates of the 92 *E. coli* ORF-*yfp* fusions during amino acid rich growth is compared against different bioinformatic measures in Fig. S9.

Codon Adaptation Index was calculated for each *E. coli* ORF using the method in (22). This calculation was implemented using the CodonAdaptationIndex class in the CodonUsage module of BioPython (version 1.58).

tRNA Adaptation Index was calculated for each *E. coli* ORF using the method in (2). This calculation was implemented using the codonR package (http://people.cryst.bbk.ac.uk/~fdosr01/tAI/index.html, downloaded on 3rd Sep 2011).

mRNA folding energy was calculated for the first 37nt of each *E. coli* ORF together with the 5 upstream nucleotides in the pZS11 plasmid backbone (GTACC). Calculation was implemented using the hybrid-ss-min command in UNAFold software (23) assuming default parameter values for reaction conditions (version 3.8, NA = RNA, T = 37, [Na+] = 1, [Mg++] = 0, maxloop = 30).

Supplementary Notes

S1. Why do we use methyl-esters in our amino acid limitation experiments?

Before we settled on the methyl-ester analog-based experiments, we tested two other amino acid limitation assays that are commonly used in the literature.

The first assay is a spin → wash → resuspend in amino acid+ / amino acid- medium (3). We did not pursue this assay for most of our experiments since it is logistically difficult to perform this

assay when working simultaneously with more than a dozen strains. However, we used this assay for the Western blotting and RT-qPCR measurements on a few strains (Fig. S1 and Fig. 4B).

The second assay involves starting with a low initial concentration of an amino acid and letting the bacterial cultures exhaust the amino acid in the medium through exponential growth(24). The bacteria then then enter the amino acid limited regime in mid-log phase without any intervention from the experimenter. However in the absence of exogenous sources of amino acid in the amino acid limited regime, protein synthesis occurs only transiently for less than an hour under these conditions and YFP synthesis rates from all *yfp* variants drop below measurable levels at the end of this time period (Fig. S15). More importantly, there is no extended steady state during which differential protein synthesis rates can be measured accurately. Nevertheless, we have confirmed that the measurements with and without methyl-esters give qualitatively similar results (Fig. S16). In addition, Western blotting done in the absence of methyl-ester reproduced the large difference in protein levels between synonymous variants of YFP during amino acid limitation (Fig. S1).

In contrast to the assay without methyl-ester, presence of methyl-ester analogs in the growth medium results in a quasi-steady state of amino acid limited growth due to hydrolysis of the ester during which differential YFP expression can be measured easily (Fig. S15). Such partial amino acid limited growth is also likely to be the relevant scenario when prototrophic strains run out of amino acids in their growth media and have a limited supply of amino acids through protein degradation or partially up-regulated biosynthesis pathways.

S2. What is the effect of varying the initial concentrations of amino acids and methyl-esters in our amino acid limitation assay?

Increasing the initial concentration of the amino acid or its methyl-ester results in a higher cell density for the onset of amino acid limitation, and when the corresponding concentrations are decreased, this onset happens at a lower cell density. Importantly, the observed differential robustness of protein synthesis (such as the data shown in Fig. 1D) is qualitatively the same upon 2-fold changes to the initial concentration of either the amino acid or its methyl-ester. As an extreme example, see Figs. S15 and S16 for comparison between the cases with and without methyl-ester analog in the growth medium.

Supplementary References

1. Dong H, Nilsson L, Kurland CG (1996) Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates. *J Mol Biol* 260:649–663.

2. dos Reis M, Savva R, Wernisch L (2004) Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res* 32:5036–5044.

3. Dittmar KA, Sorensen MA, Elf J, Ehrenberg M, Pan T (2005) Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO Rep* 6:151–157.

4. Elf J, Nilsson D, Tenson T, Ehrenberg M (2003) Selective charging of tRNA isoacceptors explains patterns of codon usage. *Science* 300:1718–22.

5. Li GW, Oh E, Weissman JS (2012) The anti-Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* 484:538–541.

6. Fiil N, Friesen JD (1968) Isolation of "relaxed" mutants of Escherichia coli. *J Bacteriol* 95:729–731.

7. Ikemura T, Dahlberg JE (1973) Small Ribonucleic Acids of Escherichia coli. *J Biol Chem* 248:5033–5041.

8. Laffler T, Gallant J (1974) spoT, a new genetic locus involved in the stringent response in E. coli. *Cell* 1:27–30.

9. Ninnemann O, Koch C, Kahmann R (1992) The E. coli fis promoter is subject to stringent control and autoregulation. *EMBO J* 11:1075.

10. Sorensen MA (2001) Charging levels of four tRNA species in Escherichia coli Rel(+) and Rel(-) strains during amino acid starvation: a simple model for the effect of ppGpp on translational accuracy. *J Mol Biol* 307:785–98.

11. Baba T et al. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2:2006.0008.

12. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *PNAS* 97:6640–5.

13. Datta S, Costantino N, Court DL (2006) A set of recombineering plasmids for gram-negative bacteria. *Gene* 379:109–15.

14. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I-1-I-2 regulatory elements. *Nucleic Acids Res* 25:1203–1210.

15. Wanner BL, Kodaira R, Neidhardt FC (1977) Physiological regulation of a decontrolled lac operon. *J Bacteriol* 130:212–222.

16. Yelverton E, Lindsley D, Yamauchi P, Gallant JA (1994) The function of a ribosomal frameshifting signal from human immunodeficiency virus-1 in Escherichia-coli. *Mol Microbiol* 11:303–313.

17. Gallant J et al. (2004) On the role of the starved codon and the takeoff site in ribosome bypassing in Escherichia coli. *J Mol Biol* 342:713–724.

18. Nagai T et al. (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol* 20:87–90.

19. Sambrook J, Russell DW (2001) *Molecular cloning: a laboratory manual* (CSHLP, Cold Spring Harbor, N.Y.). 3rd Ed.

20. Tuller T, Waldman YY, Kupiec M, Ruppin E (2010) Translation efficiency is determined by both codon bias and folding energy. *PNAS* 107:3645–50.

21. Navon S, Pilpel Y (2011) The role of codon selection in regulation of translation efficiency deduced from synthetic libraries. *Genome Biol* 12:R12.

22. Sharp PM, Li WH (1987) The Codon Adaptation Index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* 15:1281–1295.

23. Markham NR, Zuker M (2008) UNAFold: software for nucleic acid folding and hybridization. *Methods Mol Biol* 453:3–31.

24. Traxler MF et al. (2011) Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the "feast to famine" gradient in Escherichia coli. *Mol Microbiol* 79:830–845.

25. Sorensen 2005 Over Expression of a tRNALeu Isoacceptor Changes Charging Pattern of Leucine tRNAs and Reveals New Codon Reading.pdf.

26. Björk GR (1996) in *Escherichia coli and Salmonella : cellular and molecular biology* (ASM Press, Washington, D.C.).. 2nd Ed.

27. Elf J, Nilsson D, Tenson T, Ehrenberg M (2003) Selective Charging of tRNA Isoacceptors Explains Patterns of Codon Usage. *Science* 300:1718–1722.

Supplementary Tables

*Table S1: 92 ORF-*yfp *fusions used for Leu CRI validation*

Genes are arranged by increasing values of Leu CRI. $S_{Leu\text{-}rich}$ and $S_{Leu\text{-}limited}$ refer to respective protein synthesis rates (a.u. per sec per cell). Robustness refers to the ratio between the two protein synthesis rates after normalization by the corresponding value for the CTG variant of *yfp* (which is the *yfp* tag in these ORF-*yfp* fusions). ± refers to standard error of measurement.

Gene	$S_{Leu\text{-}rich}$	$S_{Leu\text{-}limited}$	Robustness	$\log_2$ Leu CRI	Gene product
polB	10.4 ± 0.3	0.6 ± 0.6	0.063 ± 0.062	-19.45	DNA polymerase II
thiH	69.7 ± 1.1	-0.7 ± 0.8	-0.012 ± 0.013	-14.52	tyrosine lyase, involved in thiamin-thiazolemoiety synthesis
aat	26.7 ± 0.8	1.3 ± 0.3	0.054 ± 0.015	-12.52	leucyl/phenylalanyl-tRNA-protein transferase
gdhA	20.4 ± 0.3	-0.2 ± 0.2	-0.013 ± 0.009	-11.62	glutamate dehydrogenase, NADP-specific
serC	91.3 ± 0.5	-1.7 ± 0.6	-0.020 ± 0.008	-11.60	3-phosphoserine/phosphohydroxythreonineaminotransferase
gpsA	111.8 ± 3.3	1.9 ± 1.9	0.020 ± 0.018	-11.38	glycerol-3-phosphate dehydrogenase (NAD+)
mlrA	44.1 ± 1.5	-0.7 ± 0.8	-0.017 ± 0.019	-11.09	DNA-binding transcriptional regulator
ybeU	41.5 ± 7.6	0.5 ± 0.3	0.017 ± 0.013	-11.00	conserved protein, DUF1266 family
argS	67.6 ± 1.9	1.0 ± 0.6	0.017 ± 0.010	-10.93	arginyl-tRNA synthetase
ilvD	59.2 ± 5.6	3.4 ± 0.9	0.064 ± 0.017	-10.39	dihydroxyacid dehydratase
ugpC	88.5 ± 2.7	1.9 ± 0.9	0.023 ± 0.011	-10.33	glycerol-3-phosphate transporter subunit
argI	44.5 ± 4.5	5.3 ± 2.6	0.138 ± 0.067	-10.24	ornithine carbamoyltransferase 1
ttdA	86.1 ± 0.2	-0.0 ± 0.5	-0.000 ± 0.006	-10.06	L-tartrate dehydratase, alpha subunit
leuS	143.4 ± 3.7	-2.5 ± 0.5	-0.019 ± 0.003	-9.46	leucyl-tRNA synthetase
hisB	119.8 ± 3.7	-0.3 ± 3.0	-0.002 ± 0.027	-9.31	fusedhistidinol-phosphatase/imidazoleglycerol-phosphatedehydratase
rpoA	71.8 ± 7.5	0.0 ± 0.6	-0.001 ± 0.009	-9.05	RNA polymerase, alpha subunit
uvrY	173.2 ± 8.8	-5.2 ± 9.3	-0.027 ± 0.061	-8.36	DNA-binding response regulator in two-componentregulatory system with BarA
rob	24.9 ± 0.7	0.5 ± 0.2	0.024 ± 0.007	-7.94	right oriC-binding transcriptional activator,AraC family
agaS	73.0 ± 3.3	1.3 ± 0.4	0.019 ± 0.005	-7.48	tagatose-6-phosphate ketose/aldose isomerase
lysS	111.8 ± 4.4	3.7 ± 2.1	0.035 ± 0.020	-7.17	lysine tRNA synthetase, constitutive
sdaB	121.0 ± 5.7	36.4 ± 4.1	0.333 ± 0.044	-7.14	L-serine deaminase II

purA	35.5 ± 2.2	-1.2 ± 0.4	-0.039 ± 0.015	-7.12	adenylosuccinate synthetase
rpoD	77.5 ± 2.2	2.3 ± 1.1	0.032 ± 0.014	-6.99	RNA polymerase, sigma 70 (sigma D) factor
melR	44.5 ± 3.3	15.1 ± 0.4	0.378 ± 0.031	-6.76	DNA-binding transcriptional dual regulator
kefF	99.2 ± 3.0	3.3 ± 0.5	0.038 ± 0.007	-6.52	flavoprotein subunit for the KefC potassiumefflux system
uhpA	22.4 ± 1.5	0.4 ± 0.7	0.019 ± 0.034	-6.50	DNA-binding response regulator in two-componentregulatory system wtih UhpB
ydcN	41.6 ± 2.7	1.9 ± 0.7	0.050 ± 0.015	-6.50	predicted DNA-binding transcriptional regulator
aspS	62.3 ± 4.2	10.9 ± 1.5	0.191 ± 0.014	-6.42	aspartyl-tRNA synthetase
relB	87.9 ± 2.3	0.8 ± 0.9	0.009 ± 0.010	-6.09	Qin prophage; bifunctional antitoxin of theRelE-RelB toxin-antitoxin system/ transcriptionalrepressor
guaA	122.6 ± 2.3	6.9 ± 2.2	0.061 ± 0.019	-5.81	GMP synthetase (glutamine aminotransferase)
phnM	203.7 ± 9.5	20.9 ± 1.2	0.113 ± 0.007	-5.81	carbon-phosphorus lyase complex subunit
tdcD	4.5 ± 0.4	7.1 ± 0.2	1.755 ± 0.163	-5.76	propionate kinase/acetate kinase C, anaerobic
ompR	114.2 ± 18.7	-1.4 ± 2.2	-0.019 ± 0.027	-5.74	DNA-binding response regulator in two-componentregulatory system with EnvZ
phnL	75.7 ± 5.3	5.7 ± 1.1	0.082 ± 0.013	-5.51	carbon-phosphorus lyase complex subunit
purH	29.6 ± 1.4	10.2 ± 0.4	0.383 ± 0.025	-5.27	fused IMPcyclohydrolase/phosphoribosylaminoimidazolecarboxamideformyltransferase
argG	2.5 ± 0.2	3.0 ± 0.2	1.343 ± 0.150	-5.07	argininosuccinate synthetase
rimM	107.5 ± 3.7	13.3 ± 0.9	0.137 ± 0.014	-4.95	16S rRNA processing protein
ubiC	93.6 ± 3.9	7.5 ± 2.2	0.087 ± 0.022	-4.87	chorismate--pyruvate lyase
leuL	48.9 ± 2.3	0.7 ± 0.5	0.017 ± 0.011	-4.75	leu operon leader peptide
asnS	9.0 ± 0.4	7.2 ± 0.5	0.884 ± 0.071	-4.65	asparaginyl tRNA synthetase
ribB	107.6 ± 7.8	12.2 ± 1.0	0.127 ± 0.020	-4.49	3,4-dihydroxy-2-butanone-4-phosphate synthase
smpB	15.8 ± 0.8	1.5 ± 0.4	0.101 ± 0.023	-4.39	trans-translation protein
guaB	31.0 ± 0.3	5.2 ± 0.6	0.184 ± 0.022	-4.21	IMP dehydrogenase
proC	69.2 ± 3.1	9.4 ± 1.6	0.151 ± 0.030	-3.92	pyrroline-5-carboxylate reductase,NAD(P)-binding
pth	89.0 ± 2.6	25.2 ± 0.9	0.313 ± 0.018	-3.10	peptidyl-tRNA hydrolase
ivbL	1.8 ± 0.3	1.6 ± 0.1	1.085 ± 0.295	-2.89	ilvB operon leader peptide
chbR	25.6 ± 1.0	24.4 ± 0.8	1.054 ± 0.008	-2.81	rRepressor, chb operon forN,N'-diacetylchitobiose utilization

leuA	78.1 ± 2.2	53.6 ± 0.8	0.756 ± 0.019	-2.77	2-isopropylmalate synthase
argD	85.7 ± 4.4	51.0 ± 1.8	0.657 ± 0.031	-2.72	bifunctional acetylornithine aminotransferase/succinyldiaminopimelate aminotransferase
yfcN	98.8 ± 2.1	24.2 ± 1.5	0.270 ± 0.011	-2.71	conserved protein
ygiD	7.3 ± 0.3	9.9 ± 0.4	1.503 ± 0.046	-2.62	predicted dioxygenase, LigB family
mdtJ	12.5 ± 1.6	17.5 ± 0.2	1.595 ± 0.204	-2.52	multidrug efflux system transporter
agaR	80.6 ± 2.5	31.4 ± 2.0	0.428 ± 0.018	-2.39	DNA-binding transcriptional repressor of the aga regulon
hisA	22.0 ± 0.9	27.6 ± 0.9	1.392 ± 0.087	-1.86	N-(5'-phospho-L-ribosyl-formimino)-5-amino-1-(5'-phosphoribosyl)-4-imidazolecarboxamide isomerase
ygbF	72.7 ± 21.5	26.4 ± 1.3	0.475 ± 0.130	-1.65	probable ssRNA endonuclease, CRISP-associatedprotein
rpoH	89.2 ± 1.8	50.5 ± 0.9	0.623 ± 0.005	-1.53	RNA polymerase, sigma 32 (sigma H) factor
leuD	128.4 ± 2.7	70.0 ± 0.6	0.601 ± 0.018	-1.48	3-isopropylmalate dehydratase small subunit
dinJ	58.9 ± 2.8	34.4 ± 1.5	0.648 ± 0.050	-1.48	antitoxin of YafQ-DinJ toxin-antitoxin system
nuoI	25.2 ± 1.8	30.6 ± 1.4	1.361 ± 0.164	-1.48	NADH:ubiquinone oxidoreductase, chain I
luxS	109.5 ± 0.7	70.3 ± 3.9	0.706 ± 0.035	-1.34	S-ribosylhomocysteine lyase
leuC	104.5 ± 6.0	54.0 ± 3.4	0.573 ± 0.069	-1.18	3-isopropylmalate dehydratase large subunit
pspA	69.8 ± 2.6	51.4 ± 0.6	0.813 ± 0.039	-1.18	regulatory protein for phage-shock-proteinoperon
pyrI	138.9 ± 13.7	112.5 ± 6.2	0.906 ± 0.091	-1.15	aspartate carbamoyltransferase, regulatorysubunit
btuE	149.1 ± 4.6	132.9 ± 3.6	0.982 ± 0.020	-0.95	glutathione peroxidase
msrB	153.6 ± 14.7	131.7 ± 2.8	0.958 ± 0.074	-0.80	methionine sulfoxide reductase B
coaD	94.4 ± 0.3	96.2 ± 0.8	1.122 ± 0.012	-0.71	pantetheine-phosphate adenylyltransferase
sfsB	72.2 ± 2.4	18.9 ± 1.5	0.291 ± 0.032	-0.61	DNA-binding transcriptional activator of maltosemetabolism
nirD	113.2 ± 17.8	22.1 ± 1.5	0.223 ± 0.027	-0.52	nitrite reductase, NAD(P)H-binding, smallsubunit
fdnI	62.4 ± 1.6	57.8 ± 3.4	1.019 ± 0.034	-0.43	formate dehydrogenase-N, cytochrome B556 (gamma)subunit, nitrate-inducible
greA	172.5 ± 12.4	90.0 ± 1.2	0.581 ± 0.045	-0.38	transcript cleavage factor
hupB	287.5 ± 19.6	116.5 ± 4.0	0.450 ± 0.030	-0.33	HU, DNA-binding transcriptional regulator, betasubunit
glpE	144.2 ± 2.0	142.0 ± 2.6	1.086 ± 0.035	-0.33	thiosulfate:cyanide sulfurtransferase(rhodanese)
ogrK	95.8 ± 9.5	145.2 ± 3.3	1.699 ± 0.153	-0.33	positive regulator of P2 growth

					(insertion of P2ogr gene into the chromosome)
rplD	20.9 ± 1.5	41.0 ± 1.6	2.191 ± 0.226	-0.33	50S ribosomal subunit protein L4
dmsB	116.2 ± 5.3	156.3 ± 5.1	1.486 ± 0.067	-0.28	dimethyl sulfoxide reductase, anaerobic, subunitB
rpsP	241.7 ± 3.7	118.5 ± 0.3	0.540 ± 0.007	-0.19	30S ribosomal subunit protein S16
rplX	150.6 ± 3.2	83.7 ± 3.5	0.613 ± 0.031	-0.19	50S ribosomal subunit protein L24
tpiA	138.5 ± 11.4	87.1 ± 2.2	0.700 ± 0.047	-0.19	triosephosphate isomerase
gapA	39.5 ± 2.0	62.6 ± 1.6	1.757 ± 0.135	-0.19	glyceraldehyde-3-phosphate dehydrogenase A
rpsT	174.7 ± 12.0	133.7 ± 4.5	0.849 ± 0.050	-0.14	30S ribosomal subunit protein S20
rpsJ	170.2 ± 37.2	142.3 ± 1.6	1.015 ± 0.219	-0.14	30S ribosomal subunit protein S10
rplT	120.5 ± 8.5	112.6 ± 3.7	1.039 ± 0.077	-0.14	50S ribosomal subunit protein L20
ahpC	57.0 ± 5.0	64.9 ± 1.7	1.267 ± 0.078	-0.14	alkyl hydroperoxide reductase, C22 subunit
rpsK	109.9 ± 9.2	139.0 ± 3.5	1.411 ± 0.113	-0.14	30S ribosomal subunit protein S11
tsf	200.4 ± 2.8	103.8 ± 2.1	0.570 ± 0.008	0.00	protein chain elongation factor EF-Ts
rplU	173.3 ± 12.4	92.5 ± 4.9	0.595 ± 0.056	0.00	50S ribosomal subunit protein L21
rpmI	158.7 ± 4.0	104.2 ± 1.4	0.725 ± 0.025	0.00	50S ribosomal subunit protein L35
yjgF	131.9 ± 5.3	98.1 ± 3.0	0.824 ± 0.057	0.00	conserved protein, UPF0131 family
ppiB	133.6 ± 4.5	114.6 ± 2.7	0.945 ± 0.011	0.00	peptidyl-prolyl cis-trans isomerase B (rotamaseB)
yjbJ	134.6 ± 4.4	158.7 ± 1.1	1.300 ± 0.035	0.00	conserved protein, UPF0337 family
rpsI	27.5 ± 2.8	78.6 ± 6.0	3.198 ± 0.324	0.00	30S ribosomal subunit protein S9
rpsF	31.7 ± 1.6	97.1 ± 0.8	3.392 ± 0.183	0.00	30S ribosomal subunit protein S6

Table S2: 56 ORF*-yfp *fusions used for Arg CRI validation

Genes are arranged by increasing values of Arg CRI. $S_{Arg\text{-}rich}$ and $S_{Arg\text{-}limited}$ refer to respective protein synthesis rates (a.u. per sec per cell). Robustness refers to the ratio between the two protein synthesis rates after normalization by the corresponding value for the AGA variant of *yfp* (this AGA variant was also used as the *yfp* tag in these ORF-*yfp* fusions). ± refers to standard error of measurement.

Gene	**$S_{Arg\text{-}rich}$**	**$S_{Arg\text{-}limited}$**	**Robustness**	**$\log_2$ Arg CRI**	**Gene product**
leuS	100.5 ± 2.6	16.1 ± 1.7	0.107 ± 0.012	-10.05	leucyl-tRNA synthetase
phoR	16.4 ± 3.2	0.7 ± 0.6	0.027 ± 0.029	-8.52	sensory histidine kinase in two-componentregulatory system with PhoB
glnG	61.0 ± 3.4	10.8 ± 2.3	0.118 ± 0.022	-8.51	fused DNA-binding response regulator intwo-component regulatory system with GlnL: responseregulator/sigma54 interaction protein

asnS	15.5 ± 1.8	2.6 ± 0.1	0.114 ± 0.014	-8.38	asparaginyl tRNA synthetase
guaA	91.5 ± 4.2	11.7 ± 3.7	0.083 ± 0.024	-8.25	GMP synthetase (glutamine aminotransferase)
thiH	53.9 ± 1.1	4.4 ± 1.3	0.054 ± 0.015	-7.16	tyrosine lyase, involved in thiamin-thiazolemoiety synthesis
fruR	44.7 ± 1.5	5.3 ± 0.1	0.079 ± 0.004	-6.54	DNA-binding transcriptional dual regulator
argG	10.0 ± 0.9	2.5 ± 1.1	0.175 ± 0.088	-6.49	argininosuccinate synthetase
gpsA	81.9 ± 3.9	9.1 ± 4.2	0.073 ± 0.032	-5.81	glycerol-3-phosphate dehydrogenase (NAD+)
rpoA	60.9 ± 1.6	10.5 ± 2.1	0.114 ± 0.022	-5.36	RNA polymerase, alpha subunit
yiiD	31.2 ± 1.8	9.8 ± 2.9	0.218 ± 0.078	-5.05	predicted acetyltransferase
agaR	53.2 ± 1.8	2.6 ± 0.8	0.032 ± 0.010	-5.01	DNA-binding transcriptional repressor of the agaregulon
rimK	17.9 ± 2.6	9.8 ± 1.5	0.393 ± 0.106	-4.94	ribosomal protein S6 modification protein
rpoH	75.2 ± 2.8	9.3 ± 2.9	0.081 ± 0.023	-4.76	RNA polymerase, sigma 32 (sigma H) factor
melR	43.7 ± 0.4	11.1 ± 2.3	0.170 ± 0.036	-4.46	DNA-binding transcriptional dual regulator
tdcB	32.1 ± 2.0	10.2 ± 2.9	0.208 ± 0.051	-3.83	catabolic threonine dehydratase, PLP-dependent
serC	68.2 ± 2.3	28.9 ± 4.2	0.284 ± 0.043	-3.79	3-phosphoserine/phosphohydroxythreonineaminotransferase
rnc	65.7 ± 2.5	15.4 ± 2.6	0.157 ± 0.028	-3.74	RNase III
chbR	22.3 ± 1.2	6.2 ± 2.1	0.190 ± 0.064	-3.66	rRepressor, chb operon forN,N'-diacetylchitobiose utilization
rsuA	100.3 ± 5.9	12.8 ± 2.8	0.084 ± 0.015	-3.65	16S rRNA U516 pseudouridine synthase
tauC	70.1 ± 3.2	7.4 ± 1.3	0.071 ± 0.016	-3.5	taurine transporter subunit
smpB	55.3 ± 1.3	17.5 ± 0.8	0.211 ± 0.010	-3.42	trans-translation protein
carA	28.2 ± 1.1	14.0 ± 0.8	0.331 ± 0.005	-3.4	carbamoyl phosphate synthetase small subunit,glutamine amidotransferase
ubiC	70.6 ± 1.8	4.7 ± 1.9	0.044 ± 0.018	-3.39	chorismate--pyruvate lyase
pth	71.2 ± 2.4	20.0 ± 0.8	0.187 ± 0.001	-3.21	peptidyl-tRNA hydrolase
rpsF	26.6 ± 1.9	14.7 ± 2.3	0.366 ± 0.043	-3.06	30S ribosomal subunit protein S6
gapA	34.3 ± 0.9	12.1 ± 1.9	0.234 ± 0.030	-2.86	glyceraldehyde-3-phosphate dehydrogenase A
yihL	43.2 ± 3.4	6.7 ± 1.1	0.105 ± 0.022	-2.85	predicted DNA-binding transcriptional regulator
allR	36.0 ± 1.5	17.1 ± 0.8	0.318 ± 0.026	-2.82	DNA-binding transcriptional repressor for all(allantoin) and gcl (glyoxylate) operons;glyoxylate-induced
yfcN	74.2 ± 4.8	38.9 ± 7.3	0.353 ± 0.068	-2.8	conserved protein

fdnI	42.7 ± 1.8	17.8 ± 3.2	0.277 ± 0.042	-2.65	formate dehydrogenase-N, cytochrome B556 (gamma)subunit, nitrate-inducible
ruvA	53.2 ± 4.1	26.7 ± 1.1	0.341 ± 0.040	-2.59	component of RuvABC resolvasome, regulatorysubunit
adiY	47.5 ± 8.2	6.1 ± 0.9	0.086 ± 0.003	-2.55	DNA-binding transcriptional activator
holD	15.1 ± 3.6	2.9 ± 1.3	0.153 ± 0.066	-2.43	DNA polymerase III, psi subunit
dmsB	68.1 ± 2.2	54.7 ± 6.5	0.537 ± 0.067	-2.42	dimethyl sulfoxide reductase, anaerobic, subunitB
bglJ	10.5 ± 1.3	6.6 ± 0.7	0.420 ± 0.010	-2.39	DNA-binding transcriptional activator for silentbgl operon, requires the bglJ4 allele to function; LuxRfamily
yfdT	71.4 ± 1.4	17.6 ± 0.7	0.164 ± 0.005	-2.32	CPS-53 (KpLE1) prophage; predicted protein
argF	64.9 ± 2.8	35.7 ± 2.5	0.371 ± 0.044	-2.01	ornithine carbamoyltransferase 2, chain F; CP4-6prophage
rplU	102.0 ± 7.4	20.5 ± 4.2	0.134 ± 0.029	-1.89	50S ribosomal subunit protein L21
luxS	90.3 ± 1.4	54.9 ± 2.4	0.406 ± 0.018	-1.77	S-ribosylhomocysteine lyase
coaD	70.9 ± 1.5	45.3 ± 3.4	0.426 ± 0.036	-1.75	pantetheine-phosphate adenylyltransferase
glnB	199.0 ± 6.0	61.9 ± 7.2	0.207 ± 0.021	-1.75	regulatory protein P-II for glutaminesynthetase
uidR	32.2 ± 1.7	35.7 ± 1.5	0.745 ± 0.064	-1.5	DNA-binding transcriptional repressor
relB	69.6 ± 4.3	29.8 ± 0.5	0.287 ± 0.016	-1.35	Qin prophage; bifunctional antitoxin of theRelE-RelB toxin-antitoxin system/ transcriptionalrepressor
btuE	106.1 ± 1.4	52.7 ± 4.0	0.331 ± 0.021	-1.35	glutathione peroxidase
argR	37.3 ± 3.4	13.7 ± 4.4	0.235 ± 0.054	-1.19	DNA-binding transcriptional dual regulator,L-arginine-binding
ogrK	53.3 ± 1.5	29.3 ± 0.9	0.367 ± 0.013	-1.14	positive regulator of P2 growth (insertion of P2ogr gene into the chromosome)
hupB	127.5 ± 3.6	71.2 ± 1.0	0.373 ± 0.008	-1.13	HU, DNA-binding transcriptional regulator, betasubunit
ppiB	95.0 ± 2.7	57.2 ± 10.5	0.401 ± 0.069	-1.13	peptidyl-prolyl cis-trans isomerase B (rotamaseB)
yjgF	92.2 ± 4.7	74.4 ± 3.6	0.540 ± 0.032	-1.13	conserved protein, UPF0131 family
kefF	68.6 ± 2.5	69.7 ± 7.6	0.677 ± 0.058	-1.11	flavoprotein subunit for the KefC potassiumefflux system
yjbJ	92.0 ± 5.1	96.6 ± 30.0	0.730 ± 0.272	-1.02	conserved protein, UPF0337 family
leuL	36.7 ± 1.4	31.8 ± 2.2	0.577 ± 0.023	-0.97	leu operon leader peptide
rimM	86.4 ± 1.7	89.9 ± 3.3	0.694 ± 0.029	-0.82	16S rRNA processing protein
ydjO	9.9 ± 0.9	20.7 ± 1.6	1.432 ± 0.219	-0.74	predicted protein
mdtJ	21.0 ± 0.4	26.4 ± 1.5	0.841 ± 0.060	-0.41	multidrug efflux system transporter

Table S3: 21 ORF-*yfp *fusions co-expressed with ^{GAG}Leu2 tRNA.

Genes are arranged by increasing values of Leu CRI as calculated for ^{GAG}Leu2 tRNA co-expression (see Supplmenetary methods). $S_{Leu\text{-}rich}$ and $S_{Leu\text{-}limited}$ refer to respective protein synthesis rates (a.u. per sec per cell) under ^{GAG}Leu2 tRNA co-expression. Robustness refers to the ratio between the two protein synthesis rates after normalization by the corresponding value for the CTG variant of *yfp* (see Fig. 1D). ± refers to standard error of measurement. Refer to Table S1 for corresponding values without ^{GAG}Leu2 tRNA co-expression.

Gene	$S_{Leu\text{-}rich}$	$S_{Leu\text{-}limited}$	Robustness	$\log_2$ Leu CRI	Gene product
ygiD	8.2 ± 0.6	3.8 ± 0.8	0.515 ± 0.103	-5.69	predicted dioxygenase, LigB family
chbR	26.5 ± 2.5	7.3 ± 1.5	0.299 ± 0.043	-5.68	rRepressor, chb operon forN,N'-diacetylchitobiose utilization
yfcN	92.7 ± 2.6	12.1 ± 2.9	0.142 ± 0.030	-5.47	conserved protein
mdtJ	12.7 ± 0.8	5.7 ± 1.7	0.513 ± 0.165	-4.46	multidrug efflux system transporter
ilvD	62.5 ± 1.5	23.6 ± 0.8	0.417 ± 0.022	-4.29	dihydroxyacid dehydratase
aspS	54.1 ± 2.0	18.5 ± 1.8	0.379 ± 0.046	-4.24	aspartyl-tRNA synthetase
lysS	106.3 ± 4.7	35.2 ± 0.8	0.367 ± 0.025	-4.22	lysine tRNA synthetase, constitutive
leuC	100.7 ± 0.5	19.0 ± 1.7	0.208 ± 0.019	-4.04	3-isopropylmalate dehydratase large subunit
ygbF	69.4 ± 3.4	8.1 ± 0.4	0.129 ± 0.009	-3.99	probable ssRNA endonuclease, CRISP-associatedprotein
pspA	65.1 ± 1.2	25.9 ± 2.0	0.437 ± 0.025	-3.91	regulatory protein for phage-shock-proteinoperon
guaA	122.0 ± 2.8	51.1 ± 1.9	0.462 ± 0.018	-3.27	GMP synthetase (glutamine aminotransferase)
btuE	137.7 ± 4.6	57.5 ± 3.9	0.462 ± 0.044	-3.23	glutathione peroxidase
purA	73.3 ± 1.8	34.8 ± 2.7	0.522 ± 0.035	-3.16	adenylosuccinate synthetase
ompR	116.4 ± 1.0	50.7 ± 3.2	0.479 ± 0.027	-2.86	DNA-binding response regulator in two-componentregulatory system with EnvZ
phnM	68.2 ± 4.4	48.4 ± 1.4	0.786 ± 0.044	-2.72	carbon-phosphorus lyase complex subunit
msrB	113.4 ± 3.1	37.9 ± 2.1	0.370 ± 0.028	-2.68	methionine sulfoxide reductase B
purH	71.4 ± 1.3	7.9 ± 2.0	0.120 ± 0.028	-2.61	fused IMPcyclohydrolase/phosphoribosylaminoimidazolecarboxamideformyltransferase
coaD	98.3 ± 7.4	46.7 ± 1.9	0.526 ± 0.018	-2.58	pantetheine-phosphate adenylyltransferase
rimM	114.3 ± 5.9	51.3 ± 1.1	0.498 ± 0.030	-1.99	16S rRNA processing protein
relB	79.9 ± 3.0	66.6 ± 3.3	0.923 ± 0.077	-1.73	Qin prophage; bifunctional antitoxin of theRelE-RelB toxin-antitoxin system/ transcriptionalrepressor
proC	78.1 ± 2.0	36.7 ± 4.4	0.519 ± 0.070	-1.72	pyrroline-5-carboxylate

					reductase,NAD(P)-binding

Table S4: 63 synonymous mutants of 13 different ORF-yfp fusions.

$S_{Leu\text{-}rich}$ and $S_{Leu\text{-}limited}$ refer to respective protein synthesis rates (a.u. per sec per cell) of mutant ORFs. Robustness refers to the ratio between the two protein synthesis rates after normalization by the corresponding value for the CTG variant of *yfp* (see Fig. 1D). ± refers to standard error of measurement. Refer to Table S1 for corresponding values of wild-type ORFs. The DNA sequence of the mutant variants below is provided in the gene_sequences.fasta file.

Gene	**$S_{Leu\text{-}rich}$**	**$S_{Leu\text{-}limited}$**	**Robustness**	**$\log_2$ Leu CRI**
btuE-yfp mutant 1	137.0 ± 5.3	50.1 ± 1.1	0.403 ± 0.007	-2.22
btuE-yfp mutant 4	131.5 ± 3.5	20.5 ± 0.5	0.172 ± 0.009	-3.06
btuE-yfp mutant 5	138.9 ± 1.8	34.1 ± 0.1	0.270 ± 0.003	-2.62
btuE-yfp mutant 7	123.4 ± 4.9	48.2 ± 1.8	0.433 ± 0.032	-2.22
btuE-yfp mutant 8	130.7 ± 5.2	28.8 ± 0.9	0.243 ± 0.015	-3.06
btuE-yfp mutant 9	128.9 ± 0.9	40.0 ± 0.6	0.342 ± 0.003	-2.62
chbR-yfp mutant 1	29.5 ± 1.7	17.7 ± 0.3	0.663 ± 0.028	-2.81
chbR-yfp mutant 2	26.7 ± 2.2	15.1 ± 0.4	0.634 ± 0.072	-3.25
chbR-yfp mutant 3	31.7 ± 0.9	11.3 ± 0.6	0.393 ± 0.029	-5.31
chbR-yfp mutant 4	36.0 ± 2.5	9.0 ± 0.3	0.279 ± 0.027	-6.2
chbR-yfp mutant 5	28.9 ± 2.4	10.7 ± 0.2	0.414 ± 0.033	-5.76
chbR-yfp mutant 6	32.6 ± 2.7	5.7 ± 0.5	0.197 ± 0.021	-7.03
chbR-yfp mutant 7	32.1 ± 1.2	7.7 ± 0.5	0.267 ± 0.023	-5.8
chbR-yfp mutant 9	31.3 ± 0.9	7.9 ± 0.5	0.278 ± 0.026	-6.64
coaD-yfp mutant 1	92.3 ± 1.3	39.1 ± 1.3	0.467 ± 0.021	-1.89
coaD-yfp mutant 2	91.7 ± 4.1	55.5 ± 0.8	0.669 ± 0.034	-1.1
coaD-yfp mutant 3	94.0 ± 2.4	47.7 ± 2.3	0.558 ± 0.014	-1.5
leuA-yfp mutant 6	70.5 ± 2.1	11.3 ± 0.3	0.177 ± 0.010	-4.4
leuA-yfp mutant 8	77.9 ± 2.2	6.6 ± 1.4	0.093 ± 0.020	-5.23
leuA-yfp mutant 9	70.0 ± 2.1	8.4 ± 0.2	0.132 ± 0.007	-4.84
leuB-yfp mutant 7	273.5 ± 8.7	6.1 ± 0.9	0.025 ± 0.003	-4.81
leuB-yfp mutant 8	273.3 ± 10.6	9.6 ± 1.0	0.039 ± 0.005	-4.02
leuB-yfp mutant 9	274.9 ± 8.1	8.4 ± 1.7	0.034 ± 0.007	-3.97
leuC-yfp mutant 5	86.8 ± 0.6	8.5 ± 1.7	0.108 ± 0.021	-3.64
leuC-yfp mutant 7	95.6 ± 1.9	5.1 ± 1.7	0.059 ± 0.019	-3.3
leuC-yfp mutant 8	100.5 ± 1.7	4.3 ± 0.9	0.047 ± 0.009	-3.69
leuC-yfp mutant 9	80.0 ± 4.9	14.3 ± 0.8	0.198 ± 0.013	-3.25
leuD-yfp mutant 2	123.6 ± 0.9	18.3 ± 1.2	0.163 ± 0.011	-2.76
leuD-yfp mutant 3	123.1 ± 3.7	39.5 ± 1.0	0.355 ± 0.020	-1.48
leuD-yfp mutant 4	122.0 ± 1.0	32.9 ± 0.3	0.297 ± 0.002	-1.92

mdtJ-yfp mutant 1	17.3 ± 0.6	9.1 ± 0.4	0.580 ± 0.020	-4.15
mdtJ-yfp mutant 2	12.2 ± 0.5	5.8 ± 0.3	0.520 ± 0.008	-3.31
mdtJ-yfp mutant 4	10.0 ± 0.9	4.3 ± 0.1	0.475 ± 0.047	-4.54
mdtJ-yfp mutant 5	13.6 ± 2.1	11.4 ± 0.5	0.979 ± 0.190	-4.15
mdtJ-yfp mutant 6	6.4 ± 0.9	5.1 ± 0.1	0.903 ± 0.125	-3.75
mdtJ-yfp mutant 7	5.9 ± 3.1	4.3 ± 0.3	1.251 ± 0.488	-4.15
mdtJ-yfp mutant 8	10.2 ± 1.5	8.3 ± 0.1	0.932 ± 0.117	-4.19
mdtJ-yfp mutant 9	10.3 ± 0.8	9.6 ± 0.2	1.049 ± 0.087	-3.36
msrB-yfp mutant 1	76.6 ± 1.9	23.7 ± 0.8	0.342 ± 0.014	-2.48
msrB-yfp mutant 2	65.3 ± 1.6	20.7 ± 0.8	0.348 ± 0.005	-2.92
msrB-yfp mutant 3	99.1 ± 2.7	27.8 ± 2.5	0.310 ± 0.031	-2.52
msrB-yfp mutant 4	94.0 ± 1.5	36.2 ± 0.4	0.424 ± 0.005	-2.08
pspA-yfp mutant 1	68.2 ± 1.9	16.3 ± 0.6	0.264 ± 0.006	-2.81
pspA-yfp mutant 2	64.2 ± 0.5	15.0 ± 0.7	0.257 ± 0.011	-2.85
pspA-yfp mutant 3	62.3 ± 1.5	28.5 ± 0.6	0.505 ± 0.016	-1.97
pspA-yfp mutant 5	73.3 ± 1.3	14.0 ± 0.9	0.211 ± 0.018	-3.2
pspA-yfp mutant 6	70.1 ± 3.0	18.3 ± 0.9	0.288 ± 0.013	-2.76
pspA-yfp mutant 7	73.5 ± 1.4	8.5 ± 0.2	0.127 ± 0.002	-3.64
pspA-yfp mutant 8	73.0 ± 1.2	12.3 ± 0.8	0.185 ± 0.012	-3.25
yfcN-yfp mutant 1	103.0 ± 3.5	8.3 ± 1.3	0.088 ± 0.012	-4.82
yfcN-yfp mutant 5	86.0 ± 3.6	14.9 ± 1.7	0.192 ± 0.022	-3.1
yfcN-yfp mutant 6	90.5 ± 1.6	5.6 ± 0.7	0.069 ± 0.009	-5.31
yfcN-yfp mutant 8	91.5 ± 4.1	14.1 ± 1.7	0.170 ± 0.017	-2.71
yfcN-yfp mutant 9	85.5 ± 1.1	5.2 ± 0.4	0.067 ± 0.005	-4.87
ygbF-yfp mutant 2	80.6 ± 2.6	6.2 ± 0.3	0.085 ± 0.006	-3.28
ygbF-yfp mutant 3	68.5 ± 3.9	25.4 ± 1.4	0.413 ± 0.044	-1.65
ygbF-yfp mutant 5	79.0 ± 4.1	25.1 ± 0.2	0.352 ± 0.017	-2.1
ygbF-yfp mutant 6	65.3 ± 2.9	10.8 ± 0.6	0.183 ± 0.019	-2.44
ygbF-yfp mutant 7	80.5 ± 4.0	8.3 ± 0.2	0.113 ± 0.004	-3.33
ygbF-yfp mutant 8	70.4 ± 3.1	10.8 ± 0.9	0.170 ± 0.019	-2.89
ygiD-yfp mutant 1	11.0 ± 1.0	2.1 ± 0.2	0.216 ± 0.009	-5.17
ygiD-yfp mutant 3	11.3 ± 0.9	4.2 ± 0.3	0.417 ± 0.056	-4.73
ygiD-yfp mutant 4	10.6 ± 0.5	5.4 ± 0.6	0.572 ± 0.090	-4.34
ygiD-yfp mutant 5	7.2 ± 1.1	3.3 ± 0.5	0.514 ± 0.041	-5.17
ygiD-yfp mutant 6	6.8 ± 1.2	2.7 ± 0.6	0.476 ± 0.131	-4.29
ygiD-yfp mutant 7	12.3 ± 2.5	2.7 ± 0.4	0.251 ± 0.016	-6.06

Table S5: List of strains used in this study

Limiting AA	**Strain designation**	**CGSC number**	**Genotype**

Leu, Arg	CP78	4695	W3110, argH- leuB- thr- his- thi-
Ser	JW2880-1	10234	BW25113, ΔserA
Pro	JW0233-2	8468	BW25113, ΔproA
Ile	JW3745-2	10733	BW25113, ΔilvA
Gln	JW3841-1	10775	BW25113, ΔglnA
Phe	JW2580-1	10048	BW25113, ΔpheA
-	MG1655	6300	wild-type strain with known genome sequence

Table S6: Initial concentrations of amino acids and methyl-esters used for amino acid limitation experiments

Amino Acid	Amino acid concentration in overnight cultures (μM)	Amino acid concentration for amino acid limitation experiments (μM)	Amino acid methyl-ester concentration for amino acid limitation experiments (μM)	Catalog number for amino acid methyl-ester
Leu	800	100	160	L1002 (Sigma)
Arg	800	150	160	11030 (Sigma)
Ser	10000	5000	800	412201 (Sigma)
Pro	800	0	160	287067 (Sigma)
Ile	800	100	160	I0522 (VWR)
Gln	800	400	400	68604 (Astatech)
Phe	800	50	50	P17202 (Sigma)

Table S7:* $\mathbf{w_i}$ *values for Leu codons under* ^{GAG}*Leu2 co-expression

Codon	- *log* w_i (with ^{GAG}Leu2 co-expression)
CTA	0.48
CTC	0.05
CTG	0.04
CTT	0.02
TTA	0.26
TTG	0.33

Table S8: qPCR primer sequences

Gene	Forward primer	Reverse Primer
yfp	TCATGCTGTTTCATGTGATC	AGGGTGATGCTACTTATGGC
gapA	GCTGAAGGCGAAATGAAAGG	GTACCAGGATACCAGTTTCACG
rpoD	TGAAGCGAACTTACGTCTGG	AGAACTTGTAACCACGGCG
rpoS	TCTCAACATACGCAACCTGG	AGCTTATGGGACAACTCACG

rpoH	TCGTAATTATGCGGGCTATGG	CAGTGAACGGCGAAGGAG
rpoE	CCAGAAGGGAGATCAGAAAGC	TACCACATCGGGAACATCAC
rpoN	TGATCCAACTCTCCCAATTCG	TCGTGATTGGCTAACAGATCG
fecI	ACTACCACAGCTTCCTTAACG	TTTCGCTGACCATTACCCG
fliA	CGCTATGCTGGATGAACTTCG	CTAAACGTTCCGCTACCTCAG
leuA	GTCGCTAACTACAACGGTCG	GCACGCCAGATATTGTTCAG
ilvM	GTTTCCACGTCTGCTCAATG	CTGACTAAACAGTAAGTCGACCG
ilvB	TGAGTTTCCGTGTCCAATCC	ATCTGATGCTGACCAACGTC

Table S9: Codon- tRNA assignments

tRNA	**unmodified anticodon**	**modified anticodon (if known)**	**cognate codons**	**reference**
Leu1	CAG	-	CUG	(1)
Leu2	GAG	-	CUC, CUU	(1)
Leu3	UAG	cmo^5UAG	CUA, CUG, CUU	(1, 25)
Leu4	CAA	CmAA	UUA, UUG	(1, 26)
Leu5	UAA	$cmnm^5UmAA$	UUA, UUG	(1, 26)
Arg2	ACG	ICG	CGU, CGC, CGA	(1, 26)
Arg3	CCG	-	CGG	(1)
Arg4	UCU	mnm^5UCU	AGA, AGG	(1, 26)
Arg5	CCU	-	AGG	(1, 27)
Ser1	UGA	cmo^5UGA	UCA, UCU, UCG	(1, 26)
Ser2	CGA	-	UCG	(1)
Ser3	GCU	-	AGC, AGU	(1)
Ser5	GGA	-	UCC, UCU	(1)
Pro1	CGG	-	CCG	(1)
Pro2	GGG	-	CCC, CCU	(1)

Pro3	UGG	cmo^5UGG	CCA, CCU, CCG	(1, 26)
Ile1	GAU	-	AUC, AUU	(1)
Ile2	CAU	k^2CAU	AUA	(1, 26)
Gln1	UUG	mnm^5s^2UUG	CAA, CAG	(1, 26)
Gln2	CUG	-	CAG	(1)
Phe	GAA	-	UUC, UUU	(1)